



# Estimation of Multiple Inflows and Effective Channel by Assimilation of Multi-satellite Hydraulic Signatures: The Ungauged Anabranching Negro River

Léo Pujol, Pierre-André Garambois, Pascal Finaud-Guyot, Jerome Monnier, Kévin Larnier, Robert Mose, Sylvain Biancamaria, Hervé Yesou, Daniel Medeiros Moreira, Adrien Paris, et al.

## ► To cite this version:

Léo Pujol, Pierre-André Garambois, Pascal Finaud-Guyot, Jerome Monnier, Kévin Larnier, et al.. Estimation of Multiple Inflows and Effective Channel by Assimilation of Multi-satellite Hydraulic Signatures: The Ungauged Anabranching Negro River. Journal of Hydrology, 2020, pp.125331. 10.1016/j.jhydrol.2020.125331 . hal-02925863v2

**HAL Id: hal-02925863**

**<https://hal.science/hal-02925863v2>**

Submitted on 18 Sep 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimation of Multiple Inflows and Effective Channel by Assimilation of Multi-satellite Hydraulic Signatures: The Ungauged Anabranching Negro River

L. Pujol(1)\*, P-A. Garambois(2)(1)(3), P. Finaud-Guyot(1)(4)(5)(6), J. Monnier (7)(8), K. Larnier(9)(7)(8), R. Mosé (1)(6), S. Biancamaria(10)(11), H. Yesou(14)(1), D. Moreira (12), A. Paris (13), S. Calmant(10)(11)

(1) *Laboratoire des Sciences de l'ingénieur, de l'informatique et de l'imagerie (ICUBE), Fluid Mechanics Team, CNRS, Université de Strasbourg, France*

(2) *Now at: INRAE (Irstea), Aix Marseille Univ, RECOVER, Aix-en-Provence, France*

(3) *INSA Strasbourg, Strasbourg, France*

(4) *Now at: HSM, Univ Montpellier, CNRS, IRD, Montpellier, France*

(5) *Now at: LEMON, INRIA, Montpellier, France*

(6) *ENGES, Strasbourg, France*

(7) *Institut de Mathématiques de Toulouse (IMT), France*

(8) *INSA Toulouse, France*

(9) *CS corporation, Business Unit Espace, Toulouse, France*

(10) *Laboratoire D'Etudes En Geophysique et Oceanographie Spatiales (LEGOS, UMR 5566 CNES CNRS IRD UPS)*

(11) *Université de Toulouse III Paul Sabatier, OMP, Toulouse, France*

(12) *UFRJ/CPRM, Av. Pasteur 404, 22290-040 Rio de Janeiro, Brazil*

(13) *Collecte Localisation Satellite (CLS), Toulouse, France*

(14) *SERVICE Régional de Traitement d'Image et de Télédétection (SERTIT), Université de Strasbourg, Pole API, Illkirch, France*

---

## Abstract

With the upcoming SWOT satellite mission, which should provide spatially dense river surface elevation, width and slope observations globally, comes the opportunity to assimilate such data into hydrodynamic models, from the reach scale to the hydrographic network scale. Based on the HiVDI (Hierarchical Variational Discharge Inversion) modeling strategy (Larnier et al. [1]), this study tackles the forward and inverse modeling capabilities of distributed channel parameters and multiple inflows (in the 1D Saint-Venant model) from multisatellite observations of river surface. It is shown on synthetic cases that the estimation of both inflows and distributed channel parameters (bathymetry-friction) is achievable with a minimum spatial observability between inflows as long as their hydraulic signature is sampled. Next, a real case is studied: 871 km of the Negro river (Amazon basin) including complex multichannel reaches, 21 tributaries and backwater controls from major confluences. An effective modeling approach is proposed using (i) WS elevations from ENVISAT data and dense in situ GPS flow lines (Moreira [2]), (ii) average river top widths from optical imagery (Pekel et al. [3]), (iii) upstream and lateral flows from the MGB large-scale hydrological model (Paiva et al. [4]). The calibrated effective hydraulic model closely fits satellite altimetry observations and presents real like spatial variabilities; flood wave propagation and water surface observation frequential features are analyzed with identifiability maps following Brisset et al. [5]. Synthetic SWOT observations are generated from the simulated flowlines and allow to infer model parameters (436 effective bathymetry points, 17 friction

patches and 22 upstream and lateral hydrographs) given hydraulically coherent prior parameter values. Inferences of channel parameters carried out on this fine hydraulic model applied at a large scale give satisfying results using noisy SWOT-like data at reach scale. Inferences of spatially distributed temporal parameters (lateral inflows) give satisfying results as well, with even relatively small scale hydrograph variations being inferred accurately on this long reach. This study brings insights in: (i) the hydraulic visibility of multiple inflows hydrographs signature at large scale with SWOT; (ii) the simultaneous identifiability of spatially distributed channel parameters and inflows by assimilation of satellite altimetry data; (iii) the need for prior information; (iv) the need to further tailor and scale network hydrodynamic models and assimilation methods to improve the fusion of multisource information and potential information feedback to hydrological modules in integrated chains.

*Keywords:* 1D Saint-Venant Model, Hydrology couplings, Variational Assimilation, Satellite Altimetry, SWOT, Hydraulic Visibility, Ungauged River

---

\* Corresponding author: l.pujol@unistra.fr

## 1. Introduction

Hydrographic networks represent major flowpaths for freshwater in the water cycle and an interface with the space of human societies. It is of prior importance in a context of climate change to improve the knowledge and representation of continental water fluxes, including river discharge, defined as an essential physical variable (see Global Climate Observing System [6, 7]). However, modeling flows structure in the different compartments of a catchment remains a hard task (see [8] and references therein) especially at poorly gauged locations. In complement of in situ sensors networks, which are declining in several regions (e.g. [9]), new generations of earth observation satellites and sensors provide increasingly accurate and dense measurements of water surface variabilities.

The Surface Water and Ocean Topography (SWOT) satellite, to be launched in 2021, will bring observations of water surface (WS) with an unprecedented spatio-temporal coverage [10, 11, 12, 13, 14]. This will yield greater hydraulic visibility (see definition in [15, 16, 17]) of hydrological responses through WS signatures from the local scale to the hydrographic network scale, hence an opportunity to better characterize hydrological fluxes and potentially constrain local to integrated hydrodynamic models and inverse problems. However, estimating river discharge  $Q$  from “geometric” observables of flow surface (elevation  $Z$ , width  $W$  and slope  $S$ ) remains a difficult inverse problem particularly in case of poor knowledge on river bathymetry and friction (see [18, 1] and references therein).

Hydraulic inverse problems with various model complexities, data-unknowns types and amounts are investigated by recent studies in a satellite data context (see [13] for a review). A few studies started to test the benefit of assimilating (synthetic) SWOT WS observations with sequential methods in simplified hydraulic models, for estimating inflow discharge assuming known river friction and bathymetry [19, 20] or inferring bathymetry assuming known friction [21, 22]. Next, methods based on low-complexity models have been proposed for estimating river

discharge from WS observables in case of unknown bathymetry  $b$  and friction  $K$ , based on the low Froude model [23, 18], hydraulic geometries [24] or empirical algebraic flow models [25]. The intercomparison of low complexity methods in [26] highlights the difficulty of estimating the so-called unknown triplet  $(Q, K, b)$  from WS observables as well as the importance of good prior guesses on the sought parameters.

The combined use of dynamic flow models of river systems and optimization methods enables to solve hydraulic inverse problems, as shown for upstream flood hydrograph(s) estimation by [27] from WS width time series and a 1D Saint-Venant model or by [28, 29, 30] using variational assimilation of flow depth time series in a 2D shallow-water model. The variational data assimilation (VDA) approach (see e.g. [31] and references therein) is suitable to address the present hydraulic inverse problem from WS observations (see [32, 5, 33, 34, 1] and references therein - single upstream hydrographs in all studies except multiple “stepwise” offtakes on synthetic and densely observed irrigation-like cases in [32]). It consists in fitting the modeled flow features to observations through the optimization of control parameters in a variational framework. To be solved efficiently, such an ill-posed inverse problem needs to be regularized: see [35] for the theory of regularization of such inverse problems and [32, 1] for the present inverse flow problem.

Crucial aspects of this difficult inverse problem are (i) the spatio-temporal sparsity of altimetric observations regarding flow controls – as analyzed in [5] for inferable hydrographs frequencies with the introduction of the *identifiability maps* and in [34] for inferable channel parameters patterns; (ii) the sensitivity of the triplet inference to good prior guesses on the sought parameters as highlighted in a SWOT context by [18, 36, 1, 37, 34]. The latest is highlighted by recent discharge estimates (in a triplet setup) from synthetic SWOT data on the Pô, Garonne and Sacramento Rivers in [1] (see also [38]), from AirSWOT airborne measurements on the Willamette River in [37] or from ENVISAT altimetric data on an anabranching portion of the Xingu River [34]. Using a biased prior hydrograph results in a biased estimate of inflow hydrograph despite a correct temporal variability at observation times - see [1] for detailed analysis. A hierarchical modeling strategy HiVDI (Hierarchical Variational Discharge Inversion) is proposed in [1] including low complexity flow relations (Low Froude and locally steady-state) for providing robust prior guesses to the VDA process by taking advantage of databases or regional hydrological models.

Most studies mentioned above tackle the estimation of a single upstream inflow discharge hydrograph from WS observations on relatively short river reaches regarding the spatio-temporal sparsity of (satellite) observations sampling and without complex flow zones - confluences, multichannel portions (except [34]), floodplains. Moreover, few recent studies address the effective modeling of (ungauged) river channels using multisatellite data [15, 39, 34, 40].

The present study investigates the challenging inference of multiple inflows and channel parameters patterns from hydraulic signatures in a SWOT context. Particular attention is paid to the difficult inference of hydraulic controls (HC) with correlated effects on WS signatures including overlapping backwater effects. Moreover, we present an effective hydraulic modeling approach based on multi-satellite observations of WS and accounting for hydrological



model inputs. It is applied to a long river reach including confluences with tributaries and strong backwater effects in the Amazon basin. The computational inverse method, based on the full 1D Saint-Venant equations, is that presented in [5, 1] with a spatially distributed friction power law in water depth and a simple piecewise linear channel bathymetry [34]. It is adapted here to account for lateral inflows/offtakes and is weakly coupled to the large scale MGB hydrological model [41, 42, 43]. Numerical investigations of the resulting WS signatures and identifiability tests are presented along with sensitivity analysis to the parameters of both the (forward) hydraulic model and the inverse method. The challenging inference of multiple inflows and channel parameters patterns is investigated with various observations densities including the assimilation of synthetic SWOT ones.

The paper is organized as follows. Section 2 presents the modeling approach with the 1D Saint-Venant flow model and the inverse computational method. Section 3 investigates the capabilities of the inverse method for identifying spatially distributed inflows with and without unknown channel parameters given observation patterns of WS signatures including overlapping backwater effects. Section 4 presents the effective modeling approach from multisatellite data applied to 871 km of the Negro river (Amazon basin) and the analysis of flow propagation features against SWOT observability. Section 5 proposes inference tests for spatially distributed inflows with and without unknown parameters on the Negro case in the presence of strong backwater effects.

## 2. Modeling Approach

### 2.1. The flow model

The Saint-Venant equations ([44]) consist in the unidirectional form of the shallow water equations and are commonly used to describe open channel flows (see e.g. [45, 46, 47] for detailed assumptions including the long wave one). In what follows,  $x$  denotes the curvilinear abscissae from upstream to downstream along a reach of length  $\mathcal{L}$  (usual simplifying hypothesis are used) and  $t \in [0, T]$  denotes the time. In this representation, let  $A(x, t)$  be the flow cross sectional area  $[\text{m}^2]$  and  $Q(x, t)$  the discharge  $[\text{m}^3.\text{s}^{-1}]$  such that  $U = Q/A$  represents the longitudinal cross-section averaged velocity  $[\text{m}.\text{s}^{-1}]$ . The Saint-Venant equations in  $(A, Q)$  variables at a flow cross section read as follows:

$$\begin{cases} \partial_t A + \partial_x Q & = k_{lat} q_{lat} \\ \partial_t Q + \partial_x \left( \frac{Q^2}{A} \right) + gA \partial_x Z & = -gAS_f + k_{lat} U q_{lat} \end{cases} \quad (1)$$

where  $Z(x, t)$  is the WS elevation  $[\text{m}]$  and  $Z = (b + h)$  with  $b(x)$  the river bed level  $[\text{m}]$  and  $h(x, t)$  the water depth  $[\text{m}]$ ,  $R_h(x, t) = A/P_h$  the hydraulic radius  $[\text{m}]$ ,  $P_h(x, t)$  the wetted perimeter  $[\text{m}]$ ,  $g$  is the gravity magnitude  $[\text{m}.\text{s}^{-2}]$ ,  $q_{lat}(x, t)$  is the lineic lateral discharge  $[\text{m}^2.\text{s}^{-1}]$ , and  $k_{lat}$  is a lateral discharge coefficient chosen equal to one here since we consider inflows only. In DassFlow, the friction term  $S_f$  is classically parameterized with the empirical Manning-Strickler law established for uniform flows:

$$S_f = \frac{|Q|Q}{K^2 S^2 R_h^{4/3}} \quad (2)$$

The Strickler friction coefficient  $K$  [ $\text{m}^{1/3}.\text{s}^{-1}$ ] is defined as a power law in  $h$ :

$$K(x, h(x, t)) = \alpha(x) h(x, t)^{\beta(x)} \quad (3)$$

where  $\alpha$  and  $\beta$  are spatially distributed parameters. This spatially distributed friction law enables a variation of friction effects in function of the flow state (see effective modeling of multichannel flows in [34]).

Inflow hydrographs  $Q_{in}(t)$  and  $q_{lat,l}(t)$  at  $l \in [1..L]$  are classically imposed respectively upstream of the river domain and at known injection cells, that is inbetween two computational cross-sections along the river channel. Let us recall the Froude number definition  $\text{Fr} = U/c$  comparing the average flow velocity  $U$  to pressure wave celerity  $c = \sqrt{gA/W}$  where  $W$  is the flow top width [ $m$ ]. Considering subcritical flows ( $\text{Fr} < 1$ ) in a satellite observability context (see [18]), a boundary condition is imposed at the downstream end of the model using the Manning-Strickler equation depending on the unknowns  $(A, Q, K)_{out}$ . The initial condition is set as the steady state backwater curve profile  $Z_0(x) = Z(Q_{in}(t_0), q_{lat,1..L}(t_0))$  for hot-start. This 1D Saint-Venant model (eq. (1)) is discretized using the classical implicit Preissmann scheme (see e.g. [48]) on a regular grid of spacing  $\Delta x$  using a double sweep method enabling to deal with flow regimes changes,  $\Delta t$  is precised in numerical cases. This is implemented into the computational software DassFlow [49].

## 2.2. The computational inverse method

The paper studies the estimation of spatially and temporally distributed flow controls from WS observables using the inverse method presented in [1] (see also [5]) with an augmented composite control vector  $c$ ; the method is detailed in AppendixA. The principle of the inverse method is to estimate (discrete) flow controls by minimizing the discrepancy between observed and simulated flow lines,  $Z_{obs}$  and  $Z(c)$  respectively, the latter depending on the unknown parameters vector  $c$  through the hydrodynamic model (eq. (1)). This discrepancy is classically evaluated with the observation cost function term  $j_{obs}(c) = \frac{1}{2} \|Z(c) - Z_{obs}\|_{\mathcal{O}}^2$  computed on the observation spatial and temporal grids, see details in AppendixA.

The control vector  $c$  contains temporally and spatially distributed unknown “input parameters” of the 1D Saint-Venant model: a friction law ([34]) and lateral inflows, unlike in [1], where there is a spatially uniform friction law  $K(h)$  without lateral flows. It reads:

$$c = (Q_{in}^0, \dots, Q_{in}^P; Q_{lat,1}^0, \dots, Q_{lat,1}^P, Q_{lat,2}^0, \dots, Q_{lat,L}^P; b_1, \dots, b_I; \alpha_1, \dots, \alpha_N; \beta_1, \dots, \beta_N)^T \quad (4)$$

where  $Q_{in}^p$  is the upstream discharge (the superscript  $p \in [0..P]$  denotes the observation time),  $Q_{lat,l}^p$  is the lateral discharges injected in the inflow cell  $l \in [1..L]$  (note that  $Q_{lat,l} = q_{lat,l}\Delta x$ ),  $b_i$  the river bed elevation ( $i \in [1..I]$ )

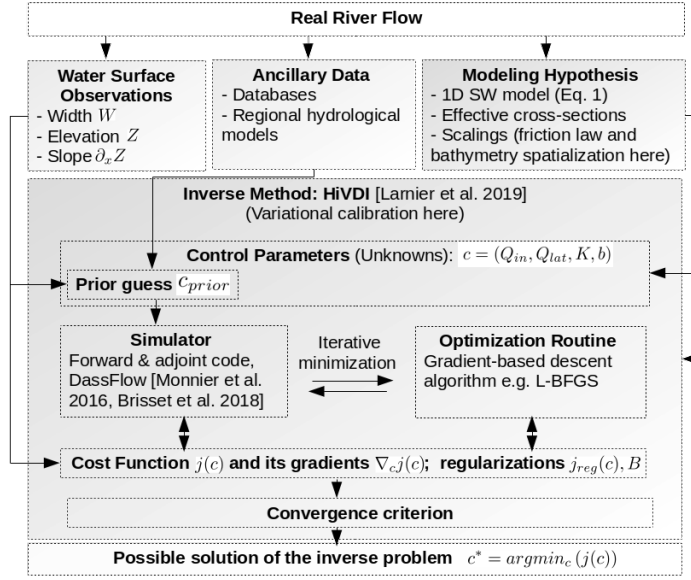


Figure 1: Flowchart of the method using the HiVDI inverse method [1] for variational calibration, adapted from [50, 51, 34].

denotes the computational cross section index in space) and, for each patch  $n \in [1..N]$  with  $N \leq I$ , the spatially distributed parameters  $\alpha_n$  and  $\beta_n$  of the friction law (eq. (3)) depending on the flow depth.

The inversion consists in solving the following minimization problem:  $c^* = \operatorname{argmin}_c j(c)$  starting from the so-called prior  $c_{prior}$  in the parameter space. This minimization problem is solved using a first order gradient-based algorithm, more precisely the classical L-BFGS quasi-Newton algorithm (see AppendixA). Note that the sought parameters have a correlated influence on the modeled flow lines, therefore leading to an ill-posed inverse problem. In order to be solved efficiently, the optimization problem is “regularized” as detailed in AppendixA. The main steps of the method are illustrated in Fig. 1.

### 3. Inference capabilities from WS signatures: synthetic test cases

In order to calibrate the parameters of a hydraulic model (eq. (1)) from WS observables, one has to identify and understand the influence of these parameters on the observable(s): in our case the WS profile. Fluvial flows are studied here in the context of satellite altimetry (see [18]). Following [52], the influence of the parameters on the modeled flow lines is referred to as their “hydraulic signature” (HC) and a reach is defined inbetween two fluvial HCs. Fluvial HCs can be defined in steady state (see [52]) as “*local maximal deviations of the flow depth from the normal depth  $h_n$  (equilibrium), imposing the upstream variation of the water depth profile  $h(x)$  over the so-called control length [53]*”. They can stem from a change in either the hydraulic resistance, cross-section shape, bottom slope or total flow variation through lateral exchanges.

This section studies the influence of inflows on hydraulic signatures, the capabilities of the inverse method described above to infer multiple inflows and channel parameters (either spatially constant or not), with a focus on

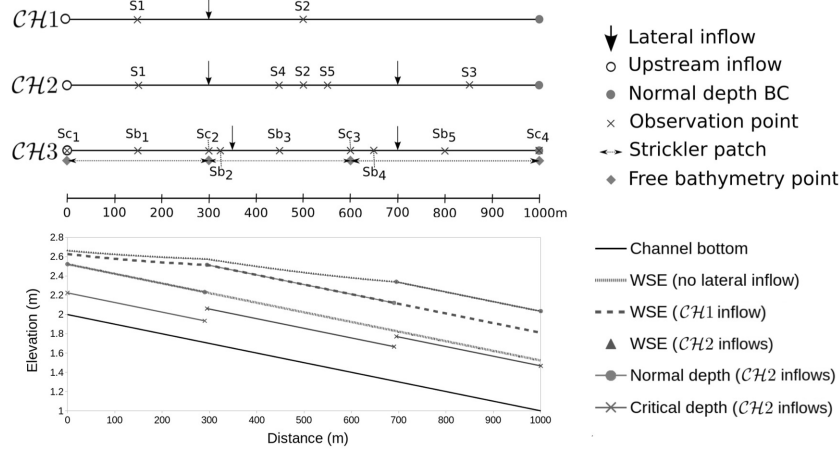


Figure 2: **TOP:** Academic test cases configurations. Rectangular channels of length  $\mathcal{L} = 1000$  m and constant cross section width  $W = 300$  m, constant bottom slope of  $10^{-3}$  m/m for *Ch1,2* and varying between  $10^{-4}$  and  $10^{-2}$  m/m for *Ch3* - the bottom  $b(x)$  is defined by linear interpolation between the 4 bathymetry points (diamonds,  $b = \{2, 1.88, 1.28, 1.12\}$  m) - and friction defined by constant values on 3 patches ( $\alpha = \{30, 12.5, 30\} \text{ m}^{1/3} \cdot \text{s}^{-1}$ ). Upstream inflow at  $x = 0$ ; for *Ch1,2,3* respectively lateral injections at abscissae (in m):  $x = 300$ ,  $x = \{300, 700\}$ ,  $x = \{350, 700\}$ , and observations at  $\{x_{S1}, \dots, x_{S5}\} = \{150, 500, 850, 450, 550\}$ ,  $\{x_{Sb1}, \dots, x_{Sb5}\} = \{150, 325, 450, 600, 800\}$ ,  $\{x_{Sc1}, \dots, x_{Sc4}\} = \{0, 300, 600, 1000\}$ . **BOTTOM:** Sample waterlines with visible upstream and downstream controls and signatures. For the sake of clarity here, upstream and injected flow are set at  $100 \text{ m}^3 \text{ s}^{-1}$  ( $\text{Fr} \sim 0.12 - 0.3$ ). Using the identifiability index  $I_{ident} = T_{wave}/\Delta t_{obs}$  introduced in [5] with  $T_{wave} = \mathcal{L}/c_k$  and the kinematic wave velocity for a rectangular channel  $c_k = 5/3U$  ( $c_k = 1.16 \text{ m} \cdot \text{s}^{-1}$  considering average speed  $U = 0.69 \text{ m} \cdot \text{s}^{-1}$ ) and a high observation frequency ( $\Delta t_{obs} = 20 \text{ s}$ ), gives a high identifiability index  $I_{ident} = 43$  for the present flow observation configuration.

the influence of the spatial observability of those hydraulic signatures.

### 3.1. Test case design

Three test cases configurations representing typical hydraulic-observations setup of increasing complexities involving lateral inflows are presented (see Fig. 2). Cases *Ch1* and *Ch2* are designed to study the effect on the inference of the overlapping signatures triggered by the propagations of, respectively, one or two lateral hydrographs, concomitantly with the one of the upstream inflow hydrograph. Case *Ch3* is a complexification of *Ch2* through the introduction of a non-flat bottom and a variable friction pattern  $K = \alpha(x)$  as needed in a real river case in the next sections ( $\beta = 0$  in eq. (3) - see investigations on spatialized friction laws with multiscale bathymetry controls in [34]).

For all three channels the boundary conditions (fluvial) consist in: (i) a normal depth (equilibrium) imposed downstream and (ii) sinusoidal hydrographs (see Tab. 1) imposed upstream and at lateral injection cells. The simulation time step is set to  $\Delta t = 20 \text{ s}$  for all cases. They are set up as twin experiments, where a forward run of the flow model (eq. (1)) is used to generate perfect WS elevation observations which are then used to infer an unknown parameter vector  $c$  (eq. (4)) with the inverse method described in section 2.2 and AppendixA. The inferences are started from erroneous prior guesses  $c^{(0)}$  that verify Manning-Strickler law for hydraulic consistency, that is unbiased priors (see investigations in [1, 34]); hydrograph priors are constant values equal to the average value of the target hydrographs.

Increasingly challenging inverse problems are considered, with increasing number of unknowns sought simultaneously and various observations densities. Cases *Ch1* and *Ch2* are used to infer temporal parameters only, given

Case	$Q_{in}$			$Q_{lat,1}$			$Q_{lat,2}$			Froude range
	$Q_0$	$a_Q$	$T$	$Q_0$	$a_Q$	$T$	$Q_0$	$a_Q$	$T$	
<i>Ch1</i>	100	0	6300	100	20	6300	-			0.13-0.29
<i>Ch2a</i>	100	20	6300	100	20	6300	100	20	6300	0.12-0.3
<i>Ch2b</i>				400	80	6300				0.05-0.55
<i>Ch2c</i>				100	20	630				0.1-0.33
<i>Ch2d</i>				100	20	6300				0.12-0.3
<i>Ch3</i>				100	20	6300				0.09-0.53

Table 1: Parameter values for sinusoidal hydrographs  $Q(x, t) = Q_0(x) + a_Q(x) \sin\left(\frac{2\pi}{T}t\right)$  used in synthetic channels; resulting modeled Froude ranges. Flows in  $\text{m}^3\text{s}^{-1}$ , time  $T$  in s.

a channel of constant slope and friction. Case *Ch3* is the most challenging case with all inflows and non constant channel parameters sought simultaneously.

### 3.2. Informative content of hydraulic signatures: single/multiple inflows inferences

The fluvial signature from a single lateral inflow is divided in two parts (see *Ch1* on Fig. 2, bottom): (i) in the reach downstream of the injection point, the cumulative flow ( $Q = Q_{in} + Q_{lat,1}$ ) is uniform with a water depth corresponding to the normal depth imposed downstream, (ii) in the reach upstream of the injection point an M1 backwater curve profile (see [54], [55, 52] in the present “altimetry context”) is obtained given the upstream flow  $Q_{in}$  and the water depth imposed downstream of this reach as the normal depth corresponding to the cumulative flow. In the case of two distinct lateral injections (*Ch2*), WS signatures overlap in the most upstream reach because of the stronger backwater effect created by two downstream inflows, which represent a more challenging inference problem.

Inference trials in case *Ch1* with control vector  $c_1 = \left(Q_{lat,1}^0, \dots, Q_{lat,1}^P\right)^T$ , assuming a known constant  $Q_{in}(x=0, t)$ , show that a single observation point in space with a dense sampling in time, placed either upstream (*S1*, in Fig. 2, top) or downstream (*S2*) from the lateral inflow, is sufficient to infer one lateral inflow hydrograph perfectly - noiseless twin experiments - (not presented). Indeed, the hydraulic signature of a lateral inflow is visible and fully informative either upstream from it because of its downstream control on the upstream flow line or downstream from it, in the signature of the cumulative flow. This means that as long the river is well temporally-observed regarding its response time (see [5] without lateral inflows) and that the temporal variations of the observed system stem from a single control, only one spatial point is needed to infer this parameter.

In the case of two distinct lateral injections (*Ch2*), WS signatures overlap in the most upstream reach because of the stronger backwater effect created downstream by the two inflows, which represents a more challenging inference problem considering the unknown control vector  $c_2 = \left(Q_{in}^0, \dots, Q_{in}^P; Q_{lat,1}^0, \dots, Q_{lat,1}^P, Q_{lat,2}^0, \dots, Q_{lat,2}^P\right)^T$ . Several variants of *Ch2* are considered to study the possible misattribution of flow controls (locations, amplitudes and frequencies) in case of identical inflow hydrographs (*Ch2a*), the backwater influence of inflow hydrographs on  $Q_{in}$  downstream signature observed at *S1* given 4 times larger inflow amplitude (*Ch2b*) or 10 times higher frequency

225 (*Ch2c*), different observations samplings “mixed” inflows signatures (see Fig. 2 and Tab. 1).

### 226 3.2.1. Inference of multiple inflows

227 For all cases, using perfect and dense observations in space (1 every 10 m) and also in time leads to quasi perfect  
228 inferences. The influence of a sparser sampling and of the observability patterns of overlapping WS signatures on  
229 the identifiability of multiple inflows with the present inverse method is studied here - without a priori weighting  
230 of the parameters in the inverse method, that is equal and unadjusted  $\sigma_{\square}$  values (see AppendixA). The inferred  
231 hydrographs are summed up in Fig. 3. Scores are given in Tab. 2, including cost function values and iterations  
232 number at convergence.

233 *Variant Ch2a.* Given only one observation station by reach (*S1*, *S2*, *S3*) very satisfying inferences of the 3 inflows  
234 are obtained (Fig. 3, red line). Hence sufficient information is provided by those three stations observing distinct  
235 signatures in each reach from upstream to downstream: (*S1*) propagation of the inflow  $Q_{in}(x = 0, t)$  in presence of  
236 the overlapping backwater effects due to  $Q_{lat,1}(x = 300, t)$  and  $Q_{lat,2}(x = 700, t)$ ; (*S2*) propagation of  $Q_{in}(x = 0, t) +$   
237  $Q_{lat,1}(x = 300, t)$  in presence of the overlapping backwater effect due to  $Q_{lat,2}(x = 700, t)$ ; (*S3*) the propagation of  
238 the total discharge without downstream control.

239 *Variant Ch2b.* Assimilation is more difficult than in *Ch2a* but inferred hydrographs (Fig. 3, red line) are still accu-  
240 rate (Tab. 2). This testifies to the ability to discriminate multiple sources of various amplitudes given observations  
241 of hydraulic signatures at higher frequency and at pertinent locations (*S1*, *S2* and *S3*).

242 Interestingly, this case highlights the expected misattribution behaviour between inflow sources as shown by an  
243 intermediate iteration (Fig. 3, orange line) and remaining to a lesser extent at convergence (red line):  $Q_{in}$  and  
244  $Q_{lat,1}$  are respectively over- and underestimated). This may be due to the relatively higher contribution of  $Q_{in}$  to  
245 the observed signature (it impacts WS elevation at *S1*, *S2* and *S3*) and consequently its contribution in the cost  
246 function (observation part).

247 Note that the final overestimation of  $Q_{in}$  in *Ch2b* is slightly greater than in *Ch2a*. This is likely due to greater WS  
248 elevation variation at *S1* caused by backwater from  $Q_{lat,1}$ , which is first attributed to  $Q_{in}$  since it has more impact  
249 on the cost function. Remember that, with perfect observations of WS signatures, at the end of the optimization  
250 process, nearly perfect hydrographs are inferred. However, the small flow misattributions during this optimization  
251 shows the difficulty of inferring multiple controls using an observation located in a strong backwater signal.

252 *Variant Ch2c.* Perfect inferences are obtained. An intermediate iteration (Fig. 3, orange line) shows that the  
253 expected misattribution of frequencies for all 3 inflows is present, though it disappears at convergence (Fig. 3,  
254 red line). This testifies to the ability to discriminate multiple sources of various frequencies given observations of  
255 hydraulic signatures at higher frequency and at pertinent locations (*S1*, *S2* and *S3*).

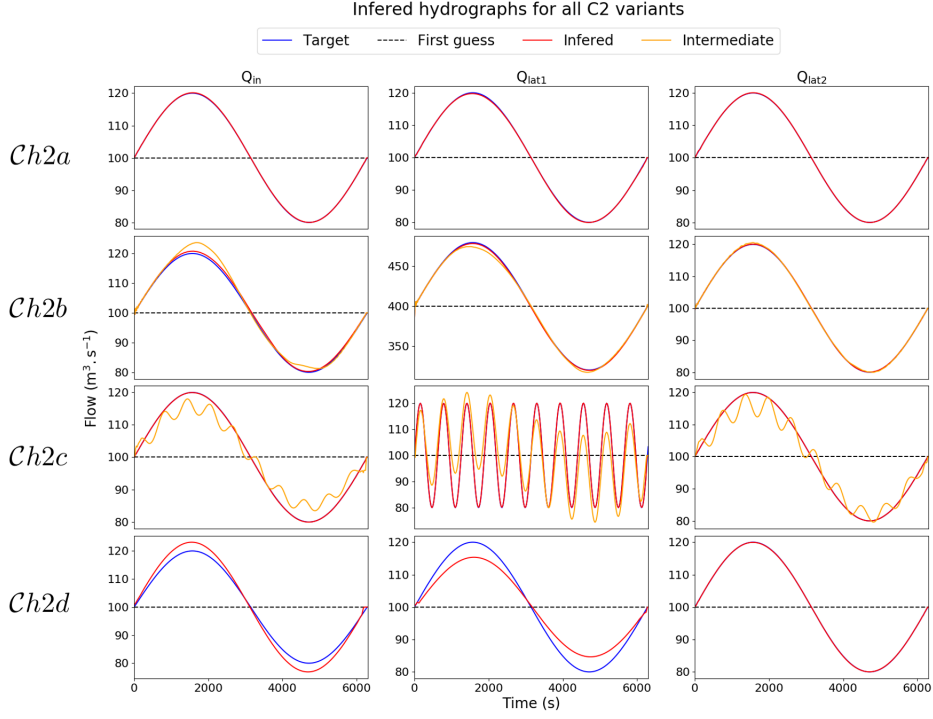


Figure 3: Inflows inferences from WS observations for all *Ch2* variants. Intermediate iteration in the assimilation process are represented for *Ch2b* and *Ch2c* ; they are hand-picked to illustrate “intermediate” behaviours before convergence (“inferred”).

256 *Variant Ch2d.* Convergence is achieved but the flow upstream of *S4* is misattributed between  $Q_{in}$  and  $Q_{lat,1}$ .  
 257 Signatures of  $Q_{in}$  and  $Q_{lat,1}$  are only observed mixed, downstream of  $Q_{lat,1}$  (at *S4* and *S5*) and downstream from  
 258 both  $Q_{lat,1}$  and  $Q_{lat,2}$  (at *S3*). Given that all stations are located in the downstream influence of both inflows, the  
 259 distribution of flow between them makes little difference on the observed WS dynamics. This confirms the need to  
 260 have at least one observation station between each sought inflow in order to be able to “separate” them.

Case	$Q_{in}$		$Q_{lat,1}$		$Q_{lat,2}$		Cost	$N_{ite}$
	RMSE	rRMSE ( $\times 10^{-6}$ )	RMSE	rRMSE ( $\times 10^{-6}$ )	RMSE	rRMSE ( $\times 10^{-6}$ )		
<i>Ch2a</i>	0.08	2.5	0.15	4.7	0.05	1.6	$8.0 \times 10^{-6}$	54
<i>Ch2b</i>	0.72	22.9	1.34	10.6	0.08	2.7	$9.4 \times 10^{-6}$	261
<i>Ch2c</i>	0.06	1.9	0.27	8.7	0.04	1.3	$7.7 \times 10^{-6}$	78
<i>Ch2d</i>	2.21	70.0	3.31	105.0	0.03	0.9	$7.9 \times 10^{-6}$	24

Table 2: Inferred parameters misfits to the truth for *Ch2* variants. The RMSE [ $\text{m}^3\text{s}^{-1}$ ] and rRMSE represent the misfit of the inferred parameters, while the cost function used in the assimilation process represents the misfit of variables.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( Q_{target}^i - Q_{inferred}^i \right)^2}, \quad rRMSE = RMSE / \sum_{i=1}^n Q_{target}^i$$

### 261 3.2.2. Synthesis

262 These first tests showed that for inferring multiple inflows, i.e. spatially distributed temporal controls, a minimal  
 263 spatial observability of their WS signature is required with one observation point between each inflow here. In case of  
 264 observation stations affected by backwater influence, the potential difficulty of separating multiple inflows from their  
 265 “mixed signature” is highlighted; using a higher spatial density of (simultaneous) observations leads to improved  
 266 inferences in the present configuration. Moreover, using observations with high temporal density (with regards to  
 267 the response time in the considered river system) and low spatial density, different frequencies can be correctly  
 268 attributed to multiple inflows (as highlighted for a single upstream inflow in [5]). Furthermore, note that if a  
 269 supercritical regime occurs in a reach between inflows, their hydraulic signatures are disconnected (not shown),  
 270 effectively reducing the assimilation problem to that of case *Ch1*.

### 271 3.3. Multiple and composite controls inference

272 In this section multiple inflows are sought simultaneously with channel parameters on case *Ch3*. Three friction  
 273 patches are consistently applied to sub-reaches inbetween the 4 sought bathymetry points. The control vector  
 274 is  $c_3 = \left( Q_{in}^0, \dots, Q_{in}^P; Q_{lat,1}^0, \dots, Q_{lat,1}^P, Q_{lat,2}^0, \dots, Q_{lat,2}^P; b_1, b_2, b_2, b_4; \alpha_1, \alpha_2, \alpha_3 \right)^T$ . Searching both inflows and  
 275 channel parameters creates a configuration (intendedly) prone to equifinality problems on the sought parameters  
 276 having correlated influence in the water surface signal. Three observation configurations (see Fig. 2) are studied:  
 277 one with a high station density (*Ch3a*: 100 stations, 1 every 10 m), another with fewer stations (*Ch3b*: 9 stations,  
 278  $Sb_{1..5}$  and  $Sc_{1..4}$ ) and a third one with even fewer stations (*Ch3c*: 4 stations,  $Sc_{1..4}$ ). Priors for inflows are those  
 279 defined for case *Ch2* (subsection 3.2.1), priors for channel parameter are hydraulically consistent with flow priors and  
 280 initial flow line. For this equifinality prone configuration, the  $\sigma_{\square}$  values used in the inverse problem regularization,  
 281 related to the sought parameters (see section AppendixA) and denoted as weights, are given in Tab. 3.

282 Inference results are presented in Fig. 4. In red, the final estimate of  $c_3$  for *Ch3a* with the “default” weights set  
 283 (see Tab. 3). In green, final inferences for variant-specific parameter weights adjusted through trial and error. In  
 284 orange, intermediate inferences with the “default” set of parameter weights. Equal values of 1, corresponding to “no  
 285 weighting”, were also tested: they lead to inaccurate inferences (not shown) and thus the “default” weights producing  
 286 more interpretable results are preferred. In further iterations, after the ones plotted in orange, behaviours similar  
 287 to the *Ch3a* “default” weights inferences (Fig. 4, in red) appear (not shown), i.e. a shift of inferred hydrographs  
 288 and Strickler coefficients away from the target. Also note that the inferred flow oscillation in the first time step  
 289 stems from the influence of the initialization scheme (see section 2.1) in the optimization on this quickly responding  
 290 channel.

291  $Q_{in}$  is underestimated while the local friction is overestimated, denoting a local tendency to equifinality. This  
 292 is linked to a strong backwater influence, created by both  $Q_{lat,1}$  and the increase in friction at  $x = 300$  m. This



Case		$\sigma_{Q_{in}} \text{ (m}^3\text{s}^{-1}\text{)}$	$\sigma_{Q_{lat,1}} \text{ (m}^3\text{s}^{-1}\text{)}$	$\sigma_{Q_{lat,2}} \text{ (m}^3\text{s}^{-1}\text{)}$	$\sigma_{\alpha} \text{ (m}^{1/3}\text{s}^{-1}\text{)}$	$\sigma_b \text{ (m)}$	Cost	$N_{ite}$
“Default”	<i>Ch3a</i>	50	50	50	600	2	0.54	180
	<i>Ch3b</i>						0.54	97
	<i>Ch3c</i>						0.63	54
“Adjusted”	<i>Ch3a</i>	1	30	30	400		0.23	156
	<i>Ch3b</i>		0.9	1.2	300		0.25	108
	<i>Ch3c</i>						0.26	100

Table 3: Parameter weight sets in *Ch3* variants.

local inflow error leads to compensation in downstream hydrographs. By adjusting parameter weights through trial and error, accurate inferences are obtained (Fig. 4, in green). This means that dense observations of the WS elevation are not sufficient for inferring all flow controls contained in  $c_3$  and that spatially distributed regularization parameters, acting as weights in the parameter search, are required.

*Variant Ch3b and Ch3c* . With sparse observations, the “default” weight set leads to worse inferences. However, the existence of a set of adjusted weights that lead to good inferences (Fig. 4, in green) is enough to show that the minimum observation spatial density of 1 station between each inflow can be sufficient to infer the extended control vector  $c_3$ . Note that adjusted weight for *Ch3b* and *Ch3c* are different from adjusted weights for *Ch3a* (see Tab. 3).

Using less observation points in space, the influence of spatial parameters decreases without loss of meaningful information and thus the relative influence of inflows increases. This simple test highlights the weighting influence of the  $\sigma_{\square}$  parameters in the regularization method in the present flow configuration. The spatial distribution and density of WS observations along with the weights change the hydraulic representativity of spatially distributed parameters in the optimization process.

The main difficulty uncovered with these academic cases is the challenge presented by simultaneous inferences of multiple inflows and channel parameters from their potentially overlapping hydraulic signatures. However, in the case of unbiased prior parameters and dense WS observation patterns relatively to those of spatio-temporal controls, satisfying inferences are obtained with the present inverse method. A real and complex river case is considered hereafter.

#### 4. Effective hydraulic model of the Negro river

After addressing increasingly challenging hydraulic inverse problems on synthetic test cases in the previous section, a real complex river flow case is now considered. It consists in 871 km of the Negro river, including several confluences with tributaries and anabranching flow zones. The reach is located upstream of the Solimoes-Negro

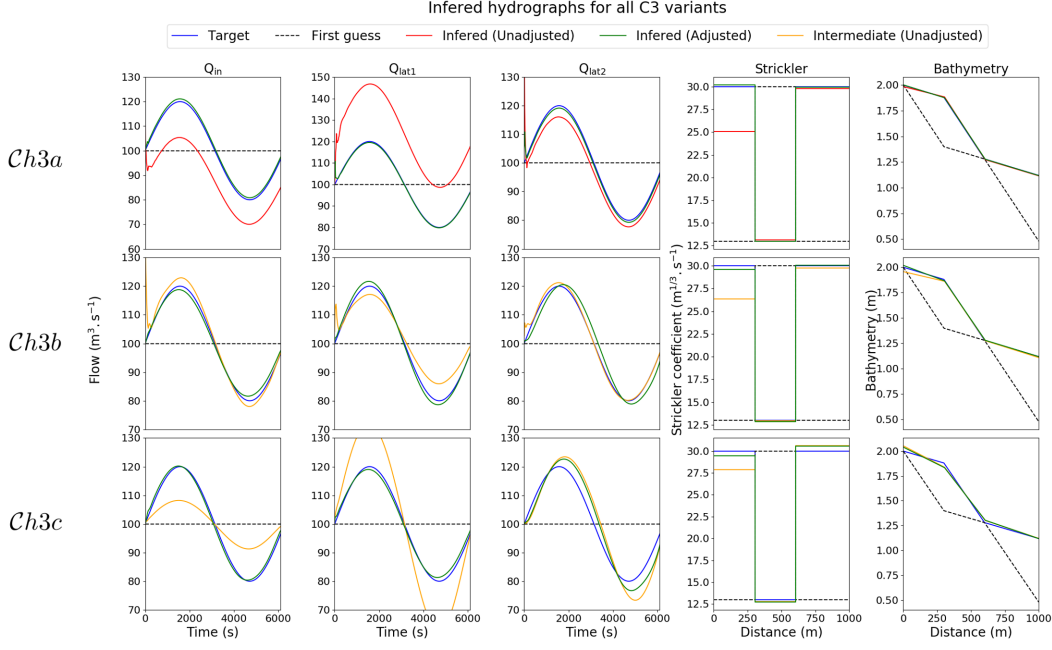


Figure 4: Inflow, bathymetry and friction patch inferences from WS observations for all *Ch3* variants. In red, final inference with “default” parameter weights (see Tab. 3). In green, final inference with adjusted parameter weights. In orange, intermediate inferences with “default” parameter weights.

confluence which is responsible for significant backwater effects (see e.g. [52]). This section presents the elaboration of effective flow models in view of performing forward and inverse flow modeling from WS observations of varying sparsity in the next section.

The modeling approach consists in (i) a 1D hydraulic model (full Saint-Venant equations, see subsection 2.1) (ii) based on effective cross sections defined from multisatellite and in situ data and (iii) weakly coupled to the large scale hydrological model MGB [41, 42, 43]. The idea is to build an effective river flow model both in coherence with the main hydrological signals (inflows) propagations along with observable flow surface signatures and hydraulic controls (see [52]). As shown in what follows, this 1D approach allows for a fair representation of flow propagation and longitudinal signatures, which are the core focus of this paper.

#### 4.1. Study zone

The study domain corresponds to the main stream of the Negro River, a major “left-bank” Amazon tributary draining the north part of the basin, with an average discharge of  $28\,400\text{ m}^3.\text{s}^{-1}$  [56]. The reach covers the 871 km upstream of its confluence with the Solimoes and presents singular channel morphologies such as multichannel flow zones mainly located in two large grabens ([57]). Part of the reach is strongly influenced by the control imposed by the Solimoes river at its confluence (average discharge of  $100\,819\text{ m}^3.\text{s}^{-1}$  according to ORE HYBAM gauge data [58], their confluence gives birth to the Amazon river). This hydraulic control is due to higher discharge and a consequently lower slope of the Negro River in its lower reach when compared to the Solimões River near to the confluence [59, 60]. The reach of interest has been crosscut by 18 ENVISAT ground tracks every 35 days from 2003

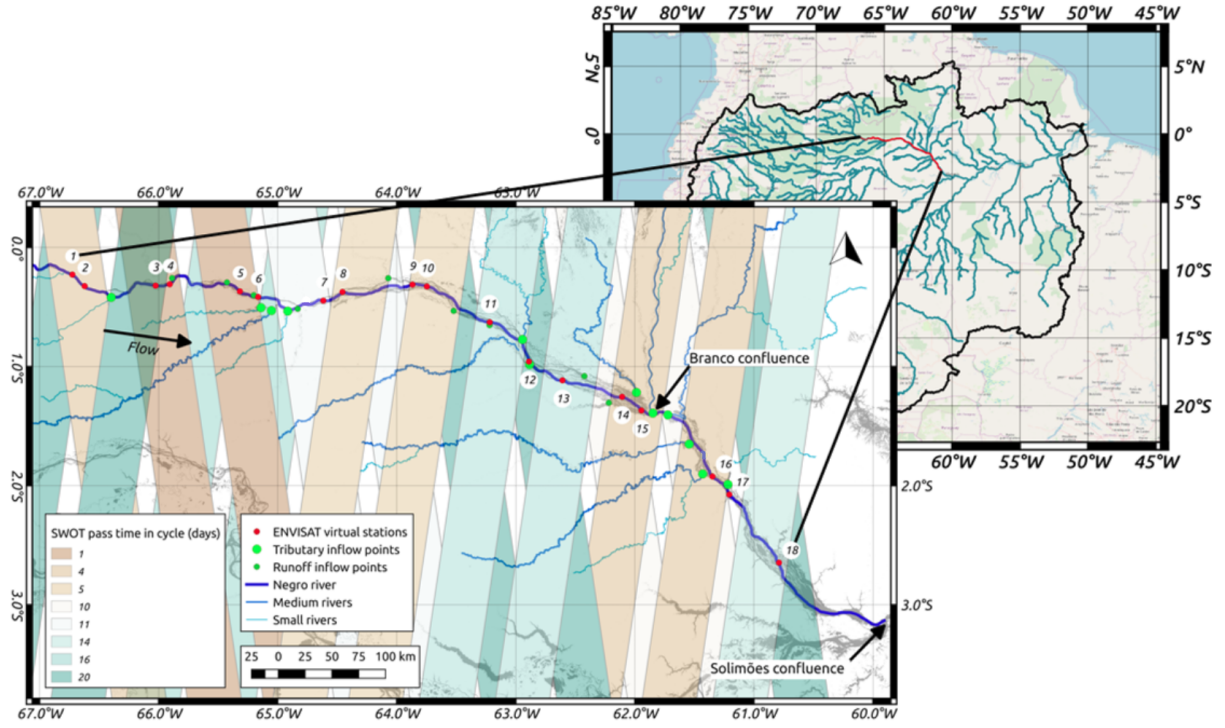


Figure 5: Study zone on the Negro river. ENVISAT Virtual Stations are numbered from 1 to 18 starting from upstream. The boundaries of the studied reach are defined by stations 1 and 18. The 21 tributaries are divided between actual rivers and inflow points from the hydrological model. SWOT swaths give an almost full spatial coverage of the reach. In gray is the average water extent map used to extract width values, from [3].

to 2010 (see [61]), representing 68 to 79 measurements of WS elevations at each of the 18 Virtual Stations (VS). Note that the measurements are not simultaneous for each station.

#### 4.2. Effective models construction

This section presents the elaboration of effective flow models from multisatellite data. First, a  $\mathcal{G}1$  “sparse” channel geometry is built from effective bottom elevations at ENVISAT VS resolution. Next, in view to test the additional constraints brought by spatially dense satellite data (synthetic SWOT), a more spatially detailed effective channel geometry  $\mathcal{G}2$  is built using a high resolution water mask and an in situ flow line as explained below.

##### 4.2.1. Effective geometry $\mathcal{G}1$ from altimetry and optical data

An effective 1D channel with effective rectangular cross sections is set up from available multisatellite data (altimetry, optical) and a large scale hydrological model following [15, 34]. According to [57], high width to depth ratios make the rectangular channel a pertinent effective modeling approach of the true geometry, even in highly anastomosed reaches - where an error on the actual hydraulic perimeter  $P_h$  hence  $R_h$  (see subsection 2.1) is expected. This is supported by a qualitative analysis of some additional ADCP measurements of river flow and cross-sectional bathymetry.

- The river centerline from [62], formed by  $30 \times 30$  m pixels, is used to calculate the river length and to project all spatial objects, such as VS, widths and inflow points, on a single one-dimensional reference.

- A longitudinal profile of cross sectional WS width  $W$  is calculated from the average river extent map derived from 31 years (1984-2015) of optic landsat imagery by [3]. A single width value per centerline point is extracted in order to build a 1D rectangular geometry. For multi-channel reaches, the effective width is the sum of the widths of all channels. This underestimates the actual hydraulic perimeter. Specific hand-filtering based on hydraulic expertise was necessary in some anabranching parts of the model where the water extent may include inactive flow zones not accounted for in the present 1D effective model. Note that [63] concurs to the necessity of reach-scale flow zone evaluation in the Amazon river catchment.
- An effective channel bottom elevation  $b_{env}$  is obtained at each VS (Fig. 6, in red) from altimetric rating curves (RC) from [64]. Its slopes range from  $-7.1 \times 10^{-5}$  to  $2.0 \times 10^{-4}$  m/m with an average of  $7.0 \times 10^{-5}$  m/m. RCs were obtained by adjusting the parameters  $(\gamma, \delta)$  of a stage discharge relationship  $Q = \gamma (Z_{sat} - b)^\delta S_{sat}^{0.5}$  using WS elevations  $Z_{sat}$  and slopes  $S_{sat}$  gained by satellite altimetry and discharge  $Q$  simulated with the large scale hydrological model MGB ([41, 42, 43]) on the temporal window of interest.

Effective rectangular cross sections geometries are defined at the  $R = 18$  VS using the above defined effective bottom elevations  $\{b_{env}\}_{r \in [1..R]}$  and river widths  $\{W_1\}_{r \in [1..R]}$ . The final model geometry ( $\mathcal{G}_1 = \{b_{env}, W_1\}_{r \in [1..R]}$ ) is obtained by linear interpolation between those 18 effective cross sections on the model grid with  $\Delta x = 200$  m.

#### 4.2.2. Effective geometry $\mathcal{G}_2$ at increased spatial resolution

Spatially dense WS elevation data is introduced in the form of an in situ GPS flow line with  $G = 579$  spatial points. It was collected by survey ship along the whole studied reach over 7 days during the low-flow period in december 2010 ([2]); it provides local WS elevations  $Z$  every 1.4 km on average and WS slopes  $S$  for every 25 km reach (ranging between  $2.0 \times 10^{-5}$  and  $8.11 \times 10^{-5}$  m/m, averaging at  $3.4 \times 10^{-5}$  m/m). Under the hypothesis of a wide rectangular cross section and a steady uniform flow, the Manning equation writes:

$$Q = K (Wh)^{5/3} (W + 2h)^{-2/3} \sqrt{S} \quad (5)$$

The water depth writes  $h = (Z - b)$  and the bottom elevation is sought using (i) the fixed WS width pattern  $W_2$  from imagery, (ii) the WS elevation  $Z_{GPS}$  and slope  $S_{GPS}$  given by the GPS profile and (iii) the discharge  $Q$  from the hydrological model (see subsection 4.1) on the river domain at the corresponding time  $t_*$ . We invert an effective bathymetry  $b_{GPS}$  using equation 5 by minimizing the square sum of misfits to  $b_{env}$  at ENVISAT stations through the modification of  $M = 14$  friction values  $((\alpha_m, \beta_m = 0), m \in [1..M], \text{friction law eq. (3)})$ . They are simply spatialized into  $M$  “hydraulic” patches consistent with large scale morphological features classified as follows: single channels, multiple channels (from 2 to 3), lightly anastomosed and heavily anastomosed (Fig. 6, in purple). The friction coefficient values are coherent with the physical properties of the classified reaches.

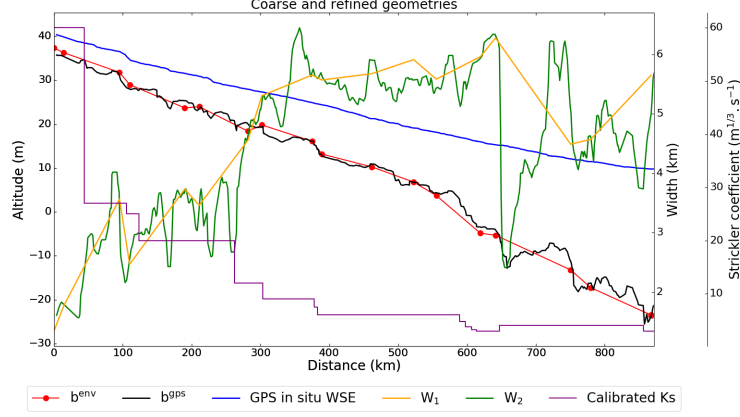


Figure 6: Effective river channel bottom and width for spatially sparse,  $\mathcal{G}_1 = \{b_{env}, W_1\}_{r \in [1..R]}$ , and spatially dense,  $\mathcal{G}_2 = \{b_{GPS}, W_2\}_{G \in [1..G]}$ , model geometries along with a low-flow GPS waterline from [2].

The new bathymetry  $b_{GPS}$  is coherent with the best available reference data and its corresponding set of physically distributed Strickler patches. The final model geometry is  $\mathcal{G}_2 = \{b_{GPS}, W_2\}_{G \in [1..G]}$ .

In the following, using either geometry  $\mathcal{G}_1$  or  $\mathcal{G}_2$ , the hydraulic model is inflowed with time series at a daily time step upstream of the river domain and at 21 tributaries (both river tributaries and runoff inflows) corresponding to the 21 catchment cells feeding into the Negro river cells in the large scale hydrological model MGB ([42, 41]). The largest of these tributaries is the Branco river at 657 km.

#### 4.3. Effective Models calibration against altimetry

The friction of the hydraulic model (eq. (1)) is calibrated against altimetric WS elevation time series following [34], i.e.  $\{Z_{s,p}^{obs}\}_{S=16, P \in [68..79]}^{env}$  at ENVISAT VS, the most downstream VS being used as BC (see subsection 2.1). The friction law is distributed using  $N = 17$  “ENVISAT” patches with constant  $(\alpha_n, \beta_n = 0)$ ,  $n \in [1..N]$  values for each reach between two successive VS. This choice is made to avoid spatial “overparameterization” in the calibration process regarding the spatial sparsity of ENVISAT observations of WS signatures. The aim of parameter calibration is to obtain a “real-like” model as close as possible of the sparse observation set. Three models are considered, to assess the impact of the bathymetry refinement and of the downstream BC on the modeled hydraulic signatures and on inverse problems: a “sparse” model ( $M1$ ) using channel geometry  $\mathcal{G}_1$  and the WS elevation time series from VS 18 as BC, a refined model ( $M2a$ ) with channel geometry  $\mathcal{G}_2$  including all the spatial variability from multisource data described above while keeping the same BC and a further changed refined model ( $M2b$ ) where the BC is changed to an altimetric RC which is of interest for “operational-like” applications in other rivers and basins.

The inverse method presented in [1] and described in subsection 2.2 and AppendixA is used here, without regularization terms, for friction calibration. Effective Strickler patches, starting from priors corresponding to average values of the “hydraulic” patches used above (Fig. 6, in purple), are calibrated following [34] who use observations of the same nature. Friction patterns  $c_{\mathcal{G}_1}^*$ ,  $c_{\mathcal{G}_2a}^*$  and  $c_{\mathcal{G}_2b}^*$  found with the inverse method are shown in Fig. 7. Most differences in calibrated friction from  $M1$  (Fig. 7, in red) to  $M2a$  (Fig. 7, in blue) correspond to their

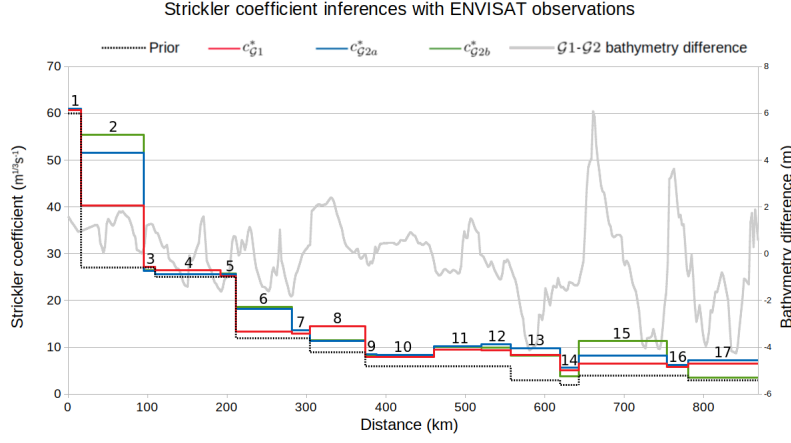


Figure 7: Friction patches after calibration against ENVISAT WS elevation observations. Inferred Strickler coefficient values are very close for all configurations for patches 1, 3 to 5, 9 to 14 and 16. Patches 2, 6 and 15 are especially sensitive to model variations.

difference in bathymetry at the virtual station point (Fig. 7, gray line), i.e. a lower slope in  $M2a$  leads to a higher inferred Strickler parameter in order to match WS observations (e.g. in patch 2 and 6. Inferred parameters for  $M2b$  roughly match those of  $M2a$ , with some discrepancies in patch 2, 15 and 17. Using a different BC influences WS sensitivity to parameters and the relative contribution to the cost function of local WS misfits, which explains differences in patch 15 and 17 ; the one in patch 2 stems for the high friction values, hence lower WS sensitivity as analyzed after.

#### 4.3.1. Water levels analysis

The following presents a detailed analysis of the effective hydraulic model for configuration  $M1$ , along with an analysis of changes obtained for configurations  $M2a$  and  $M2b$ .

The simulated WS elevation are compared to observed WS elevation at each ENVISAT virtual station in Fig. 8 - other time series are available in AppendixB. For the 3 models calibrated above, the modeled WS are fairly close to observed WS given the limited modeling complexity and data uncertainties. More precisely, the fit to the altimetric WS elevation time series is fairly good, as shown for  $M1$  in Fig. 8, and nearly unbiased as shown in Fig. 9(left). The WS elevation global RMSE is at 0.936 m for  $M1$  ; similar results are found with  $M2a$  (see Tab. 4). Errors are greater in low and high flows, with consistent underestimations of flow amplitude upstream (VS 1 – 4) which turns into overestimation downstream (VS 9 – 13), before disappearing closer to the BC (VS 14 – 18). VS 5 to 8 are particularly accurate. Error metrics are coherent with those from current state of the art models using satellite data (see e.g. [40] on the Congo river).

The analysis of the time series for  $M1$  gives insight on the 1D model behaviour regarding the real flow physics sampled with the sparse nadir altimetry data and dense in situ low flow line. Modeling errors can stem from either an (expected) improper representation of the channel and flow complexity or uncertain (ungauged) inflows and data.

Concerning the hydraulics, from downstream to upstream, relative errors are lower in anabranching reaches

425 outside of the backwater influence starting at the Branco tributary ( $x = 657$  km up to around  $x = 350$  km) and in  
 426 the backwater influence of the (known “perfect”) downstream BC. Overall, relative errors are higher upstream, in  
 427 single channelled, low water height reaches and in the Branco backwater influences. Note that 2D complex lateral  
 428 flows in floodplains or retention behaviours from “igarape” rivers may happen in high flow periods (see [65, 66]).  
 429 These unaccounted phenomenons may decrease flood wave velocities and cause hydrograph skewness ([67, 68, 10]).  
 430 The 1D modeling of water levels compared to altimetry observations (Fig. 8) can first be analyzed as follows:

- 431 • Stations 14 to 18 are located in reaches with different morphological properties. Stations 14 and 15 are located  
 432 in a densely anastomosed reach upstream of the Branco river confluence, a major tributary. Stations 16 and  
 433 17 are in single channel reaches, upstream from the confluence with the Solimoes river. Station 18 is in a  
 434 densely anastomosed reach at the location of the BC forcing on WS elevation. Their low relative misfits do  
 435 not testify to the absence of complex hydraulic behaviours in this area but rather to the dominating influence  
 436 of the BC.
- 437 • Stations 5 to 13 are located in mostly homogeneous anastomosed reaches, with stations 5 to 8 in a less  
 438 densely anastomosed region than stations 9 to 13. This spatial division corresponds to two trends in relative  
 439 misfit, where lower misfit is seen in the less anastomosed reaches. This testifies to the difficulty of modeling  
 440 potentially 2D hydrodynamics using 1D approach. Indeed, the more channels there are, the further away  
 441 the simulated wetted perimeter is from the true wetted perimeter (and so the hydraulic radius). Note that  
 442 parameterizing the Strickler coefficient as described in eq.(3) and including  $\beta(x)$  in the control vector during  
 443 the calibration process, instead of the simpler  $\beta(x) = 0$  used here, does not yield a better fit in this complex  
 444 case modeled with a single rectangular channel.
- 445 • Stations 1 to 4 are located on single channel reaches. Although the area seems the most suitable to be modeled  
 446 in 1D, it still has the highest relative misfit to ENVISAT observations. For stations 1 and 2, this is partly  
 447 due to effective width estimation errors being more prevalent in the relatively narrow channel (around 2 km  
 448 in width). Furthermore, note that effective channel bottom elevation for these stations are respectively 37.3 m  
 449 and 36.3 m while the lowest ENVISAT WS elevation observation are respectively 36.6 m and 35.8 m. This  
 450 corresponds to low-flow water heights of 0.7 m and 0.5 m which do not fit field measurements. Consequently,  
 451 relatively high friction coefficients are inferred between station pairs 1-2 and 2-3 to fit low water depth. This  
 452 misfit might be due to data error, including effective width errors for stations 3 and 4 located in areas of sharp  
 453 width variations. Note that the higher the friction values, the less sensitivity of the modeled WS elevation,  
 454 which explains the highest spread of Strickler coefficient ( $K = 40$  to  $55 \text{ m}^{1/3} \text{ s}^{-1}$ ) in reach 2 found during  
 455 calibration for the 3 models (Fig. 7).

	Global RMSE (m)	Global Average bias (m)	Upstream RMSE (m)	Upstream Average bias (m)
M1	0.94	-0.02	0.88	-0.08
M2a	0.94	-0.02	0.91	-0.01
M2b	1.72	0.23	0.90	-0.05

Table 4: RMSE and bias over 8 years for the  $M1$ ,  $M2a$  and  $M2b$  models. Upstream metrics are calculated for stations 1 to 9 only, which are outside of the BC’s backwater influence. The high global RMSE for  $M2b$  comes from the known dephasing of the Solimoes and Negro peak flow, which is not reproduced by the RC.

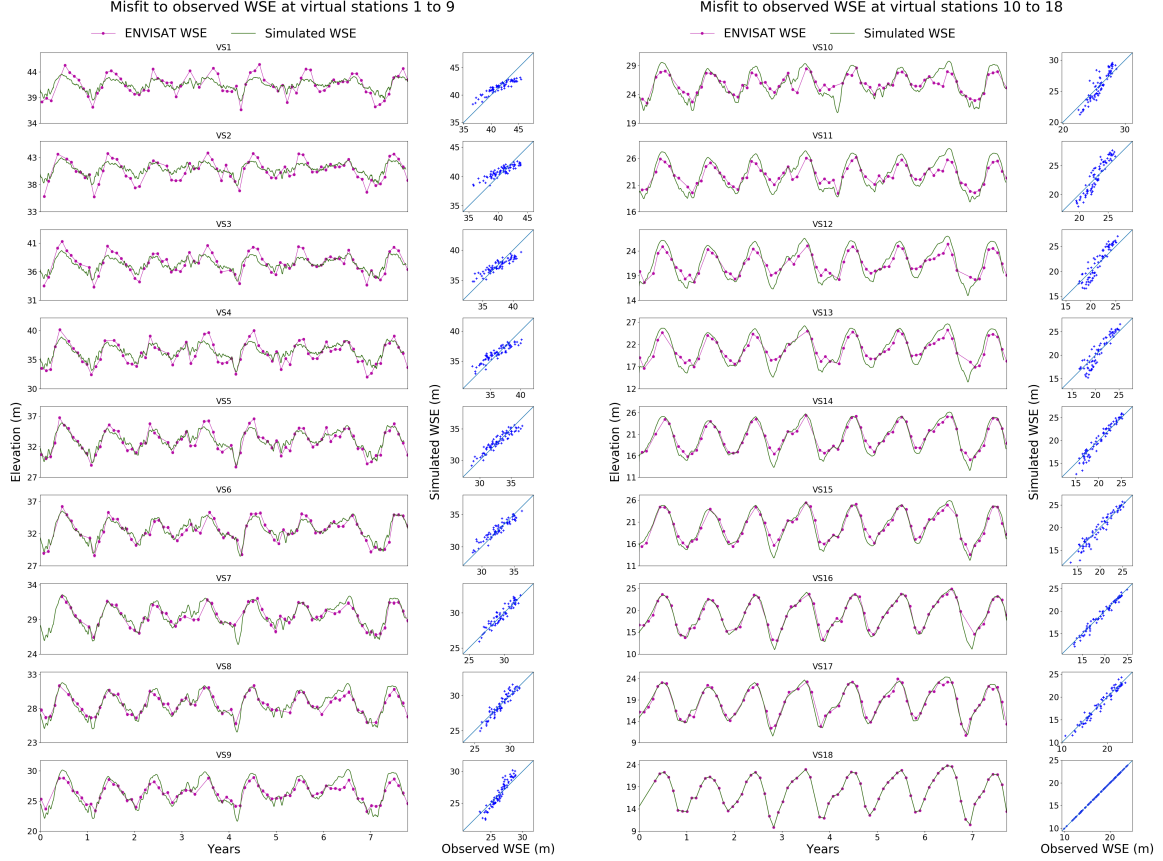


Figure 8: Modeled and observed WS elevation at ENVISAT VS after friction calibration at all stations for  $M1$

The introduction of the refined geometry  $\mathcal{G}2$  in  $M2a$ , recall for generating spatially distributed SWOT data and to perform inference tests hereafter, has low impact on WS elevation bias and errors at ENVISAT VS (see Fig. 9), with only stations 1, 2 and 3 showing significant change. Using a rating curve as downstream BC in  $M2b$  mostly impacts the downstream part of the model where some misfit to altimetry data appears. Indeed, it is more difficult, using a simple power law depending on the local flow variables, to capture the influence of the confluence with the Solimoes River - not modeled. The latest having strong discharge variations out of phase with the one of the Negro River itself (e.g. [16]).

#### 4.3.2. Effective model analysis

As a preliminary to hydraulic parameters inference from WS observables, this subsection studies the spatio-temporal features of the simulated hydraulic signatures, their sensitivity to model parameters and their observability



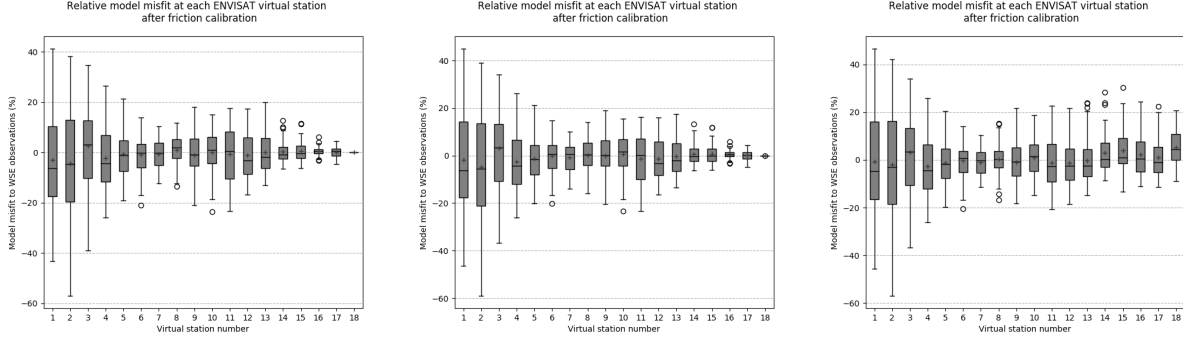


Figure 9: Relative misfit between modeled and observed WS elevation at ENVISAT for  $M1$  (left, base model)  $M2a$  (middle, WS elevation at downstream BC) and  $M2b$  (right, rating curve at downstream BC). Crosses are average values, horizontal bars are median values.

given a SWOT sampling. First, an analysis of a flood wave propagation, resulting from multiple inflows, and its hydraulic signature visibility is performed using identifiability maps following [5]. The latter consist in a space-time representation of the WS signal and flow propagation features against the observability pattern. These maps, inspired by the theory of characteristics (see [69, 47]), enable to read how the sought upstream discharge information is sampled in the downstream WS deformations and help to estimate inferable hydrograph frequencies. Next, a numerical sensitivity evaluation of the flow model is carried out.

In the context of regional hydrological modeling including river networks representation, the sensitivity of the present flow model is studied by using erroneous inputs. These inputs are also used in section 5 as erroneous priors for various assimilation setups.

- Inflow: two hydrograph sets (containing lateral inflows and the upstream BC inflow) corresponding respectively to 70% and 130% of the true hydrographs are used as erroneous values for sensitivity trials and are referred to as  $Q_{FG}^{-30}$  and  $Q_{FG}^{+30}$  respectively.
- Friction: two Strickler repartitions, with coefficient values worth respectively 70% and 130% of the truth are used as erroneous values for sensitivity and are referred to as  $K_{FG}^{-30}$  and  $K_{FG}^{+30}$  respectively.
- Bathymetry: the inflow sets  $Q_{FG}^{-30}$  and  $Q_{FG}^{+30}$  and the true Strickler values are used to dig two bathymetries as described in section 4.3. The bathymetry derived from underestimated flows is referred to as  $b_{FG}^{-30}$  (it overestimates the true bathymetry), and the other is referred to as  $b_{FG}^{+30}$ .

SWOT will provide spatially distributed observations with interesting revisit frequencies at the scales of the current river domain and hydrological signal propagations. Fig. 10 shows the evolution of the simulated WS elevation anomaly during the yearly peak flow (red-blue heatmap) as well as its multiple SWOT observability (in black). Based on the modeled flow, accounting for several inflows, the propagation of an intumescence corresponding to the annual flood wave signature is represented along the river through the maximum WS elevation in time

(following [55]) (Fig. 10, top, blue points). This intumescence propagation is visible on the upstream 400 km of the river from day 164 to day 173. It is detected by a SWOT swath at  $t = 166$  d and another one at  $t = 170$  d. It is more difficult to detect this signature in the downstream part of the river ( $x > 400$  km) affected by the strong downstream control imposed by high water depths at the Negro-Solimoes confluence; a downstream control due to the Branco tributary also overlaps from  $x = 657$  km to around  $x = 400$  km. This control can be seen through the tracked WS elevation maximum (Fig. 10, top, in gray), where an early rise in WS elevation originates from  $x = 657$  km, and through the extreme waterlines (Fig. 10, bottom, in blue), which highlights the change in length of this influence in low and high flows. As a consequence, WS observations on the downstream part may contain combined information due to the upstream hydrographs propagation but also to the expression of downstream controls.

The maximum WS elevation is tracked for simulations with erroneous parameters as defined above (Fig. 10, top, in red, green and cyan). They are not plotted where the flow displays “pool behaviour” (gray points). They highlight the sensitivity of propagation to model parameters which is also an important point when they are varied during an optimization process as featured in section 5. The propagation time from 0 to 400 km can be evaluated to around 10 days, and is estimated as follows for the rest of the river domain.

The conservative part of the Saint-Venant equations (i.e. without source terms) is hyperbolic: some quantities depending on the water depth and velocity (known as the Riemann invariants) are transported by waves at speeds different from the flow speed (see e.g. [69, 47]). The wave celerities are  $U + c$  and  $U - c$  with  $c = \sqrt{gh}$  for rectangular cross sections (see analysis of propagation features in [5]). For the fluvial regime of interest here ( $Fr = U/c < 1$ ), information propagates both downstream and upstream and the Riemann invariants are modified along the wave due to the source term effects. The wave celerities obtained on the Negro River model are given by reach in Table 5, relatively high wave speeds are obtained hence propagation of information both upstream and downstream, with spatio-temporal variability. The WS signature (and the discharge) thus reflects the nonlinear combination of information coming from both upstream (due to inflows variations) and downstream (due to local hydraulic controls or downstream BC - see the method of characteristic in [47]). This highlights the difficulty of inferring multiple inflows from sparse observations of WS signature, especially given uncertain channel parameters and backwater effects.

Nevertheless interesting frequential information can be gained from the identifiability map as introduced in the case of a single inflow. Using the kinematic wave speed  $\frac{5}{3}U$  (Fig. 10, top, dashed blue line) which compares fairly well to the intumescence speed on the upstream part of the reach (Fig. 10, top,  $x < 400$  km). This gives an approximate propagation time  $T_{wave} = 26$  d on the whole domain, greater than the SWOT observation cycle period of 21 days. This brings the reach identifiability index to  $I_{ident} = 1.23$  (defined as  $I_{ident} = T_{wave}/\Delta t_{obs}$ , i.e. the average number of time a wave is observed, see [5]). However, in the present case, the notion defined by [5] accounts for a single upstream inflow, not spatially distributed lateral inflows with potential upstream backwater controls. Actual identifiability indices for reaches in between each lateral inflow would be much lower (estimated

x (km)	0	65	121	163	193	216	388	418	465	502	528	598	623	628	657	700	739	754	789	1000
$\frac{5}{3} \bar{u}_{HF}$	2.35	1.42	1.22	1.07	1.02	0.75	0.50	0.52	0.53	0.50	0.48	0.35	0.32	0.41	0.61	0.47	0.44	0.47	0.44	
$c_{HF}$	8.32	9.23	9.71	10.31	10.36	11.19	12.52	12.61	12.62	13.49	13.95	15.90	16.28	16.82	17.93	17.26	17.56	19.37	19.69	
$u_{HF} + c_{HF}$	9.74	10.08	10.44	10.95	10.97	11.64	12.82	12.92	12.93	13.79	14.23	16.11	16.47	17.06	18.29	17.54	17.82	19.65	19.95	
$u_{HF} - c_{HF}$	-6.92	-8.38	-8.98	-9.67	-9.75	-10.73	-12.22	-12.29	-12.30	-13.19	-13.66	-15.68	-16.09	-16.57	-17.56	-16.99	-17.29	-19.09	-19.42	
$I_{ident,HF}$	0.32	0.46	0.40	0.33	0.26	2.64	0.69	1.04	0.81	0.60	1.69	0.82	0.18	0.82	0.81	0.97	0.39	0.86	2.14	
$\frac{5}{3} \bar{u}_{LF}$	1.40	0.96	0.83	0.70	0.69	0.52	0.34	0.37	0.45	0.40	0.41	0.25	0.22	0.26	0.31	0.26	0.24	0.24	0.23	
$c_{LF}$	5.54	6.16	6.54	7.13	7.05	7.72	8.70	8.59	7.87	8.71	8.81	11.30	11.83	12.54	14.11	13.37	13.78	16.09	16.59	
$u_{LF} + c_{LF}$	6.38	6.74	7.04	7.54	7.47	8.03	8.91	8.81	8.14	8.94	9.06	11.45	11.96	12.70	14.30	13.52	13.93	16.24	16.72	
$u_{LF} - c_{LF}$	-4.70	-5.58	-6.04	-6.71	-6.64	-7.40	-8.50	-8.37	-7.60	-8.47	-8.56	-11.15	-11.70	-12.39	-13.93	-13.22	-13.64	-15.95	-16.45	
$I_{ident,LF}$	0.54	0.67	0.58	0.50	0.39	3.80	1.03	1.47	0.95	0.76	1.97	1.16	0.26	1.29	1.61	1.76	0.73	1.71	4.21	

Table 5: Identifiability indexes between each pair of inflow at low and high flow (see Fig. 10, bottom). Speeds are given in  $\text{m.s}^{-1}$ .  $I_{ident}$  is given for a reach of length  $L$  and an observation time step  $\Delta t_{obs} = 1 \text{ d}$  by  $I_{ident} = \frac{L}{\frac{5}{3} \bar{u} \Delta t_{obs}}$ .

identifiabilities in between each inflow pair are given in Tab. 5 considering a fictious  $\Delta t_{obs} = 1 \text{ d}$  full domain observability). Furthermore, SWOT swaths observations consist in WS snapshots on different parts of the river domain at given times, hence containing various and mixed signatures (in the sense introduced in section 3) of both several inflows and channel parameters - the more downstream, the more aggregated is the inflow information. Inferences of multiple inflows and frequential analysis are presented in the next section given known or uncertain channel parameters, spatio-temporally dense or sparse (SWOT) observations.

## 5. Inferences from satellite observables

This section studies the challenging inference of ungauged channel parameters and multiple inflows on the Negro River case, which represents a real and complex large scale problem. Typical inverse problems in hydrological-hydraulic modeling are studied here considering SWOT WS observations. The inference of channel parameters or/and inflows in the 1D Saint-Venant model is addressed using the inverse method presented in subsection 2.2 (see also section AppendixA). The downstream BC is set as a known altimetric rating curve. Three observation sets are generated: spatially and temporally dense ( $\Delta t = 600 \text{ s}$ ) observations (DenseSet), SWOT observations from the hydraulic model outputs masked by SWOT swaths (SWOTSet) and noisy SWOT observations using the large scale simulator [70] to add realistic measurement noise (SWOTNoiseSet). We first present inferences of inflows only, then of channel parameters, and finally of all those spatio-temporal controls simultaneously.

### 5.1. Multiple hydrographs inferences

Depicting flow structure within a river network and a catchment is a key issue in hydrological modeling, especially in ungauged basins. Seeking to infer, from distributed WS observations, flow controls that are both temporally and spatially distributed can represent a very challenging inverse problem, as previously highlighted on synthetic cases.

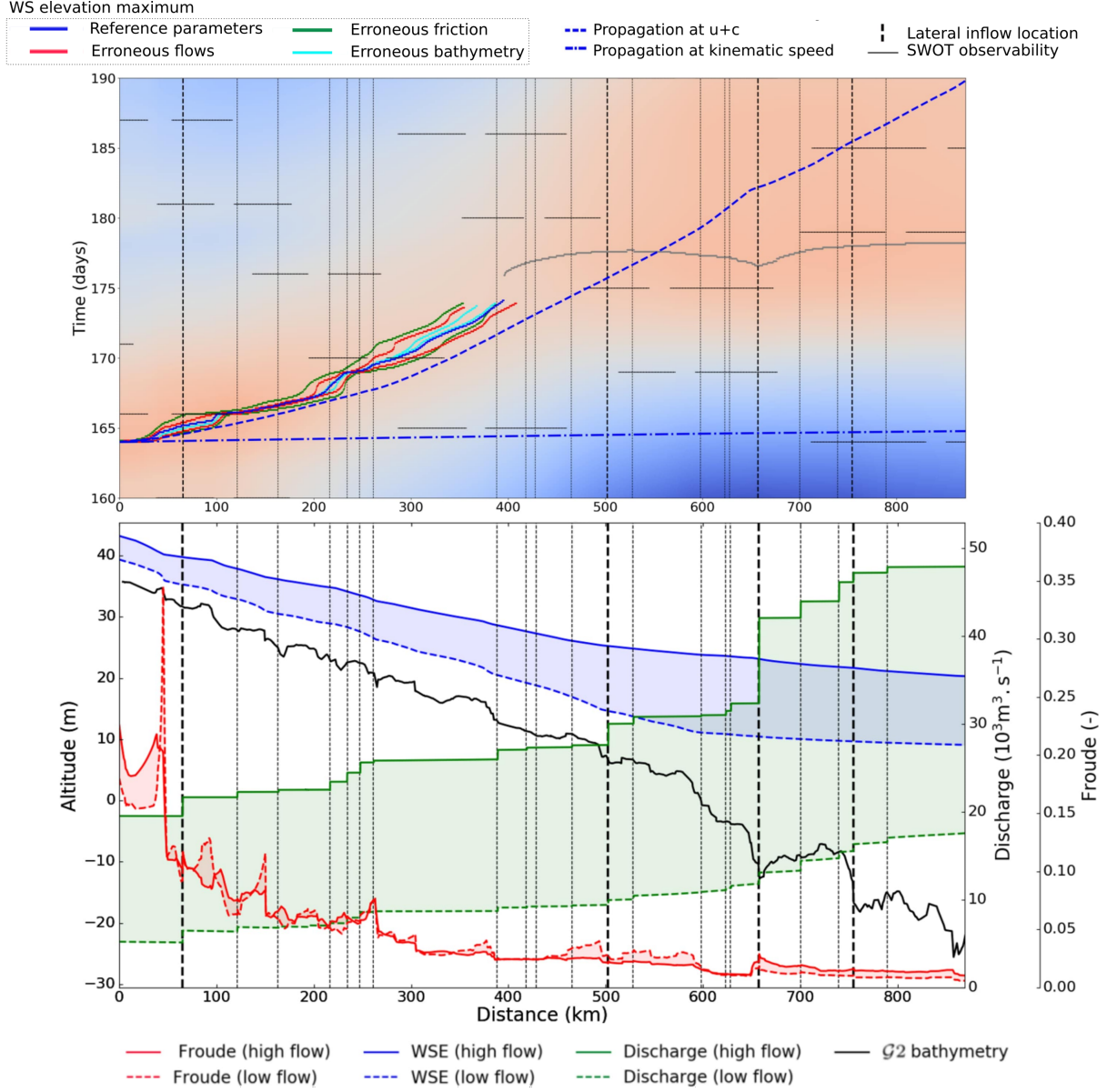


Figure 10: **TOP**: Identifiability maps and flood wave propagation during the yearly peak flow (may-june) in the Negro river model. The WS anomaly (heatmap) is given by  $Z_{ano}(x, t) = Z(x, t) - \overline{Z(x)}$ , where  $\overline{Z(x)}$  is the average local WS elevation from day 160 to 190. Blue points: tracking of maximum WS elevation value  $Z_m(x) = \max_{t \in [0, 365]} Z(x, t)$ ,  $\forall x \in [0, L]$ . Gray points: tracking of maximum WSE in the downstream pool. Dashed blue lines: fictitious trajectory at kinematic speed (sparse dashes)  $c_k = 5/3U$  and at  $U + c$  (dotted dashes), starting at  $x = 0$ , at the time of the local WS elevation peak. The speeds are calculated from the simulated flow speed  $U$  and water height  $h$  and updated every  $\Delta x = 200$  m, such as  $t^{p+1} = t^p + \frac{\Delta x}{c_k^p}$ . **BOTTOM**: Extreme flow forcings and flow model variables over a 2 year period. Blue lines: Extreme simulated waterlines. Red lines: corresponding extreme Froude values. Green lines: corresponding cumulative injected flows. Vertical black dashes are lateral inflow locations. Bold vertical dashes are inflows inferred in subsection 5.2.

544 Inferences of  $L = 21$  inflow hydrographs from 2 years of SWOT synthetic observations are studied here.  
 545 The channel geometry, friction and BCs are assumed to be known, hence the control vector reduced to  $c =$   
 546  $(Q_{lat,1}^0, \dots, Q_{lat,1}^P, Q_{lat,2}^0, \dots, Q_{lat,L}^P)^T$ . The inferences are started from a prior guess  $c^{(0)}$  consisting in true hy-  
 547 drographs affected by uncertainties of  $\pm 30\%$ , that is  $Q_{FG}^{+30}$  and  $Q_{FG}^{-30}$  as defined above. Note that the inference is  
 548 started from a hydraulically consistent initial state using an unbiased prior in the first time steps (see investigations  
 549 in [1, 34]); the prior values of regularization parameters  $\sigma_{Q_{lat}}$  correspond to inflows magnitudes.

550 The inferred hydrographs from inflow prior  $Q_{FG}^{-30}$  are presented in Fig. 11 for DenseSet (green lines) and  
 551 SWOTNoiseSet observations (orange lines). Results from prior  $Q_{FG}^{+30}$  are available in AppendixC. SWOTSet and  
 552 SWOTNoiseSet give almost identical inferences, therefore only the SWOTNoiseSet inferences are presented. For  
 553 under- and overestimated priors, the assimilation of dense and SWOT observations enables to infer the true hy-  
 554 drographs fairly well. RMSE ranges from  $8.86 \text{ m}^3 \text{ s}^{-1}$  at  $x = 465 \text{ km}$  up to  $578.31 \text{ m}^3 \text{ s}^{-1}$  for the Branco tributary  
 555 at  $x = 657 \text{ km}$ . RMSE for all inferences presented in Fig. 11 can be found in section AppendixC. Some inferences  
 556 show global under- or overestimations (e.g.  $x = 216, 388, 789 \text{ km}$ ). These biases are linked to the prior bias. Strong  
 557 and numerous overlapping backwater signals may also influence flow misattribution, as discussed in the academic  
 558 cases (section 3) for a small scale model. As tested in numerical experiments (not shown), increasing a scalar value  
 559  $\sigma_{Q_{lat,l}}$  can give more effective weight to an hydrograph  $Q_{lat,l}$  in the inference and it can be found further away from  
 560 its prior guess, which highlights the role of the covariance matrix used for regularization.

561 Note that temporal oscillations appear on the inferred hydrographs when using SWOTNoiseSet which is “tem-  
 562 porally sparse” observation patterns compared to flow propagation, which is not the case of DenseSet. These  
 563 oscillations are especially present in downstream inflows, which may link them to particular hydraulic responses  
 564 in the BC influence zone, although they can be seen in upstream inflows as well. They tend to be prevalent in  
 565 declining limbs of hydrographs (e.g. in Fig. 11, at  $789 \text{ km}$ , from day 120 to 300).

566 Note that, regardless of oscillations, inferences tend to be further from the truth in decreasing hydrographs.  
 567 These oscillations are not the effect of signal misattribution, as they are present with any number of inferred  
 568 hydrographs (not shown), nor are they caused by the prior’s shape, as filtered priors also lead to oscillations (not  
 569 shown). Instead, the oscillations seems to stem from the combination of the low observation frequency compared  
 570 to the spatially distributed inflow hydrographs and the nonlinear hydraulic response. Keep in mind that we track  
 571 flow information through WS elevation deformations caused by the nonlinear propagation of parameter signatures  
 572 (see subsection 4.3.2).

573 A sample illustration of those oscillations on the simulated WS elevation is presented in Fig. 12, at  $870 \text{ km}$ ,  
 574 downstream for three oscillating inferred hydrographs (at  $x = 738, 754$  and  $789 \text{ km}$ ). The inferred waterline from  
 575 SWOTNoiseSet is compared to the truth at all simulated times and at SWOT pass times only. The misfit is logi-  
 576 cally lowest at SWOT pass times (goal of the optimization), while unobserved periods exhibit a slightly oscillating  
 577 (unconstrained) misfit. Higher frequency observations, such as DenseSet, prevent this behaviours through a more

complete spatio-temporal observability of the WS signatures, hence constrain the spatio-temporal parameters inference further. Some model configurations where temporal parameters are discretized at a greater time step than observation one do not exhibit such behaviours (e.g. results with DenseSet, [34]). As already shown in [5] for the identification of a single hydrograph, the identification is possible only in time windows representing the wave propagation time  $T_{wave} \sim 26$  days in the present case, yet with multiple inflows and observation sample (see subsection 4.3.2). As a consequence, outside the “identifiable time windows”, the inferred values are not necessarily representing reality (see related WS misfit in Fig. 12). They are the optimal solution corresponding to the considered priors of the optimization problem. In practice, this means that introducing an additional regularization term  $j_{reg}$ , like  $\int_0^t \frac{\partial^2 h}{\partial t^2} dt$  at observation points, would smooth (as following a spline curve) between the identifiable windows instead of the obtained oscillations. This smooth discharge curve would not be more physical than the present oscillations and we made the choice to not hide this well understood phenomena. It is a logical consequence of the disparity between the samplings of observations and parameters and does not impede interpretations of hydraulic signatures and identifiability.

Seeking to infer a control that is both temporally and spatially distributed represents a challenging assimilation problem. In the present case: (i) the observation frequency now plays a role in identifying the hydraulic signature, on top of its spatial density and resulting flow propagation: (ii) varying nonlinear flow propagation, and so WS signatures, can result in different inferences depending whether they are performed from observations of rising/declining hydrographs propagations (local  $Q(Z)$  hysteresis) and (iii) indirect contributions to parameter weight in the inverse method appear, as successive hydrographs influence the whole downstream water line (established in subsection 3.2), which gives greater “effective weight” to upstream hydrographs as the cause of the misfit is observed in more stations and thus accounted for multiple times in the cost function. The inferred flow variations may be compensating for errors made at upstream stations with different SWOT pass times, impacting their WS elevation either through backwater control or through the modification of the BC and its own backwater effect.

### 5.1.1. Spatial parameters inference

The inference of effective channel parameters is studied here considering a control vector composed of all friction coefficient values and bathymetry points. The bathymetry is composed of  $I = 436$  free points (1 every 2000 m) between each of which it is obtained by linear interpolation. SWOTSet is used, with a spatial observations sampling (1 point every 200 m), i.e. 10 times greater than the sought bathymetry for observed reaches. Two inferences from hydraulically consistent priors are presented, one with the refined bathymetry  $b_{FG}^{-30}$  introduced in subsection 4.3.2 (Fig. 13, in green) and another one with  $b_{env}^{+2.6m}$ , a shifted bathymetry from the  $M1$  model in subsection 4.2.1, in red. The 2.6 m correspond to the spatially averaged shift of  $b_{FG}^{-30}$  compared to  $b_{env}$ .  $b_{env}^{+2.6m}$  does not contain a priori information on target bathymetry shape - such as a coarse DEM prior. The friction prior is  $K_{FG}^{-30}$  for both inferences.

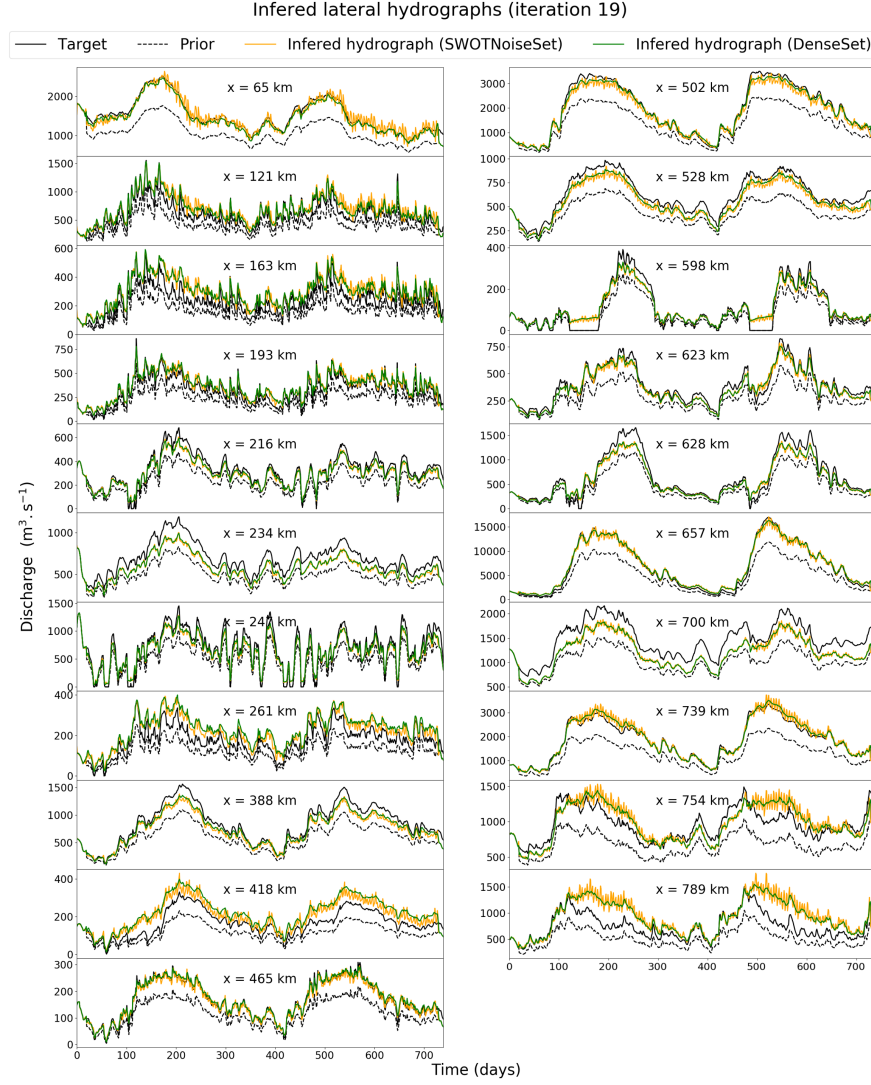


Figure 11: Lateral hydrographs inferences from SWOTNoiseSet and DenseSet, using the  $Q_{FG}^{-30}$  inflow prior .

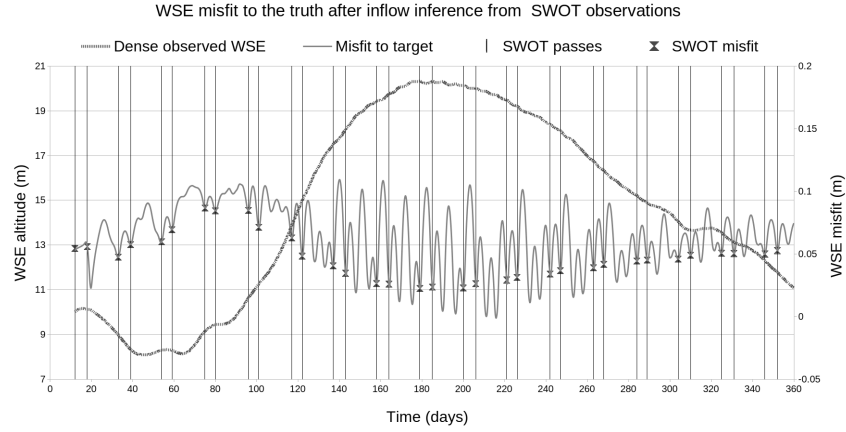


Figure 12: Difference between target and inferred WS elevation at  $870\text{ km}$ , as observed by DenseSet and SWOTSet. The dotted line represents the inferred waterline inferred from SWOTSet (with  $Q_{FG}^{-30}$  as inflow prior), but observed by DenseSet. The difference between this waterline and the target waterline is the misfit to target. At SWOT pass times, the misfit is low as expected from an inference from SWOTSet. It only displays WS elevation oscillations at unobserved times.

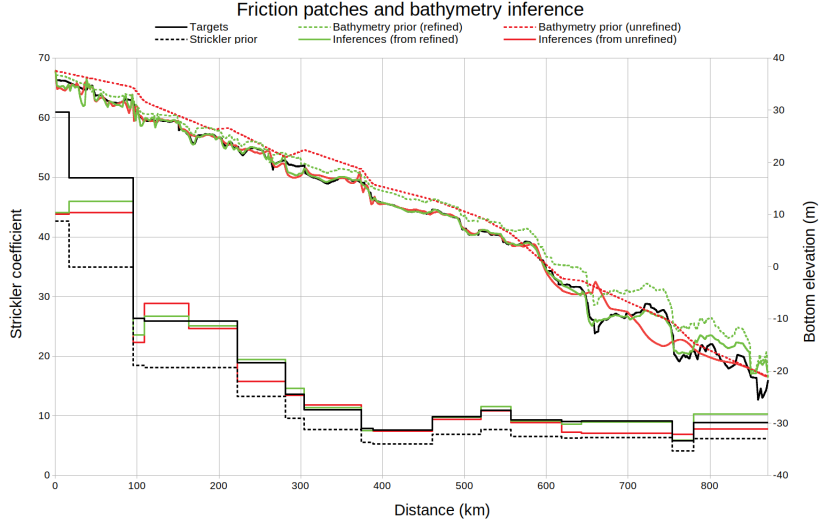


Figure 13: Bathymetry and friction pattern inferences from SWOTSet

Using  $b_{FG}^{-30}$  and  $b_{env}^{+2.6m}$  as priors, the inference of channel parameters (friction and bathymetry) respectively reach a RMSE of  $(4.362 \text{ m}^{1/3}\text{s}^{-1}, 1.231 \text{ m})$  and  $(4.762 \text{ m}^{1/3}\text{s}^{-1}, 1.913 \text{ m})$ . Upstream ( $x = 0 - 110 \text{ km}$ ), irregularities appear in both inferred bathymetries and correspond to underestimated Strickler coefficients. The high friction partially hides the hydraulic signature of the bathymetry and enables inference error inconsequential to the cost function. Downstream ( $x = 600 - 870 \text{ km}$ ), in the strong backwater control of the downstream BC, inferences tend to stay close to their prior values. Furthermore, inferences from the unrefined prior  $b_{env}^{+2.6m}$  are smoother than those from the refined prior  $b_{FG}^{-30}$ . This testifies to the role of the a priori bathymetry shape knowledge contained in the prior when the WS elevation sensitivity to bathymetry is low. Strong backwater effects make the inference of channel parameters more difficult and parameter compensation appears due to the lessened sensitivity of the simulated WS (e.g. in green, at the last friction patch). The refined bathymetry prior  $b_{FG}^{-30}$  leads to inferred bathymetry and friction that are closer to the truth. It will be used in the extended control vector inferences below.

623

## 624 5.2. Inference of channel parameters and inflows

This section investigates the simultaneous inference of both unknown inflows and channel parameters on the large scale Negro River case; it combines all previously mentioned difficulties and corresponds to an ungauged configuration. In the following, the aim is to determine: 1) if SWOT data are sufficient to infer the extended control vector given unbiased prior parameters; 2) how the added spatial complexity from lateral inflows impacts spatial parameter inference. In addition, further investigations on the impact of lateral inflow prior shape, representing for example hydrological modeling uncertainty in a simple manner, are presented. The considered extended control vector is the following:

631

$$c_{ext} = (Q_{lat,x=65}(t), Q_{lat,x=502}(t), Q_{lat,x=657}(t), Q_{lat,x=754}(t); b_1, \dots, b_H; \alpha_1, \dots, \alpha_N) \quad (6)$$



632 The inferences are performed from DenseSet and SWOTNoiseSet. The bathymetry and friction priors are  $b_{FG}^{-30}$   
633 and  $K_{FG}^{-30}$  respectively. Four major lateral inflows located at  $x = 65, 502, 657$  and  $754$  km (Fig. 10, bold dashed  
634 bars) are considered. Their reduced number facilitates the analysis of their spatial impacts. The other inflows are  
635 set to their target values. Two inflows prior types are used:  $Q_{FG}^{flat}$ , that gives no a priori on hydrograph shapes  
636 and  $Q_{FG}^{filter}$ , hydrographs obtained by applying a 80 days moving average filter to the true hydrographs. Prior  
637 flow values in  $Q_{FG}^{flat}$  are set to the target flow values from the first time step up to 120 days for the sake of initial  
638 hydraulic consistency. Inferences of all parameters for these inflow priors are presented in Fig. 14 and Fig. 15. The  
639 inferred control vectors are referred to as  $c_{flat}^*$  and  $c_{filter}^*$ . Inferred parameter scores can be found in Tab. 6.

640 Inferred spatial parameters patterns are similar to those obtained previously without unknown inflows in sub-  
641 section 5.1.1.  $c_{flat}^*$  features a fair bathymetry fit downstream ( $x = 600 - 870$  km) while  $c_{filter}^*$  stays close to the  
642 prior value. This may be due to the different range of the simulated hydraulic responses in the first iterations:  
643 using  $Q_{FG}^{flat}$  leads to an increase in WS elevation sensibility to bathymetry. Upstream ( $x = 0 - 110$  km), increased  
644 bathymetry irregularities in  $c_{flat}^*$  are linked to the erroneous prior  $Q_{FG}^{flat}$  leading to bathymetry errors in the first  
645 iterations, coupled with lower inferred Strickler coefficients, hence a lessened impact of bathymetry on the water  
646 surface and the inability to correct the “initial” errors.

647 In terms of temporal behaviours, both priors give fair estimates of hydraulic controls for DenseSet. Inferences  
648 from SWOTNoiseSet are close to those from DenseSet with the presence of oscillations and the rising part of  
649 hydrographs are better fitted than decreasing ones, as observed in subsection 5.1. In both  $c_{filter}^*$  and  $c_{flat}^*$  and  
650 for both observation sets, a correlation between sought inflows at  $x = 502; 657$  and  $754$  km appears. The Branco  
651 river flow, at  $x = 657$  km, is better inferred and its well fitted peaks are also found in the two smaller rivers (e.g.  
652 at 520 days), which are in its upstream and downstream influences zones (see Fig. 10, left). In all inferences, the  
653 total flow at the downstream BC closely matches that of the truth, which means that only hydraulic signature is  
654 misattributed across the 4 inferred temporal parameters, not on the total flow. In  $c_{filter}^*$ , more accurate inferences  
655 are obtained, with a smaller influence of the Branco river on other inflows in its influence zone and more accurate  
656 inference of small scale behaviours. The filtered prior  $Q_{FG}^{filter}$  introduces information on low frequency behaviours  
657 of the sought inflows, helping the assimilation process to converge to correct the target inflows. This configuration  
658 allows for a better fit of small scale variation in the controls.

659

## 660 6. Conclusion

661 This paper investigated the inference of river channel parameters and multiple inflows from water surface signa-  
662 tures in the context of satellite altimetry with the forthcoming SWOT mission and using water extents from optical  
663 data as well. The HiVDI inverse method presented in [1], based on the 1D Saint-Venant equations and a variational  
664 assimilation scheme adapted to account for lateral inflows (mass and momentum injections). Given hydraulically  
665 consistent prior guesses and regularization weights, it is successfully applied to synthetic test cases and a long reach

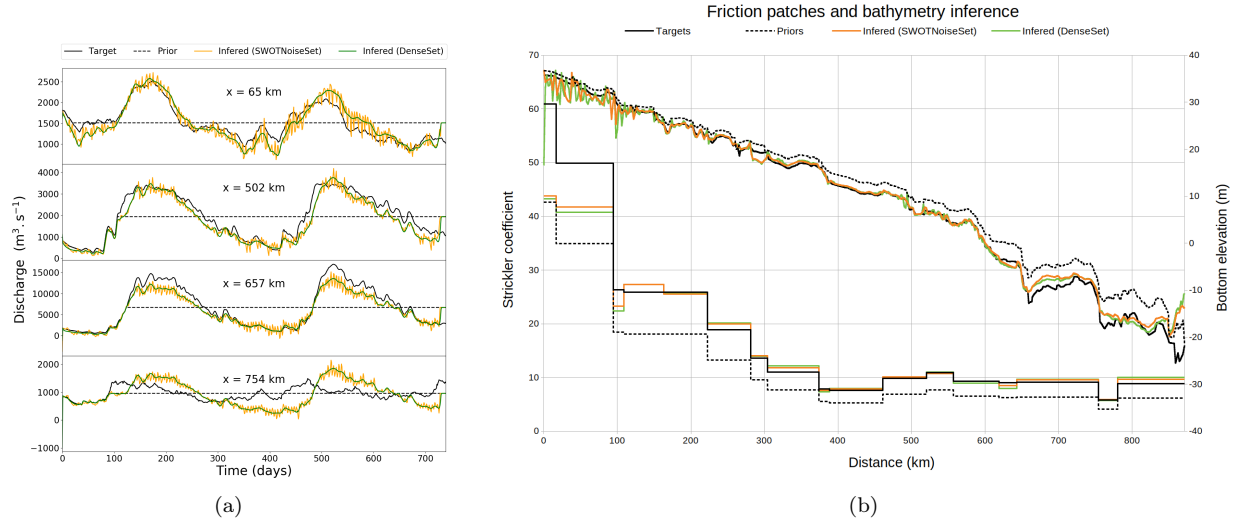


Figure 14: Inflow, bathymetry and friction patch inferences from SWOT synthetic data:  $c_{flat}^*$ , inferred control vector without a priori hydraulic behaviour.

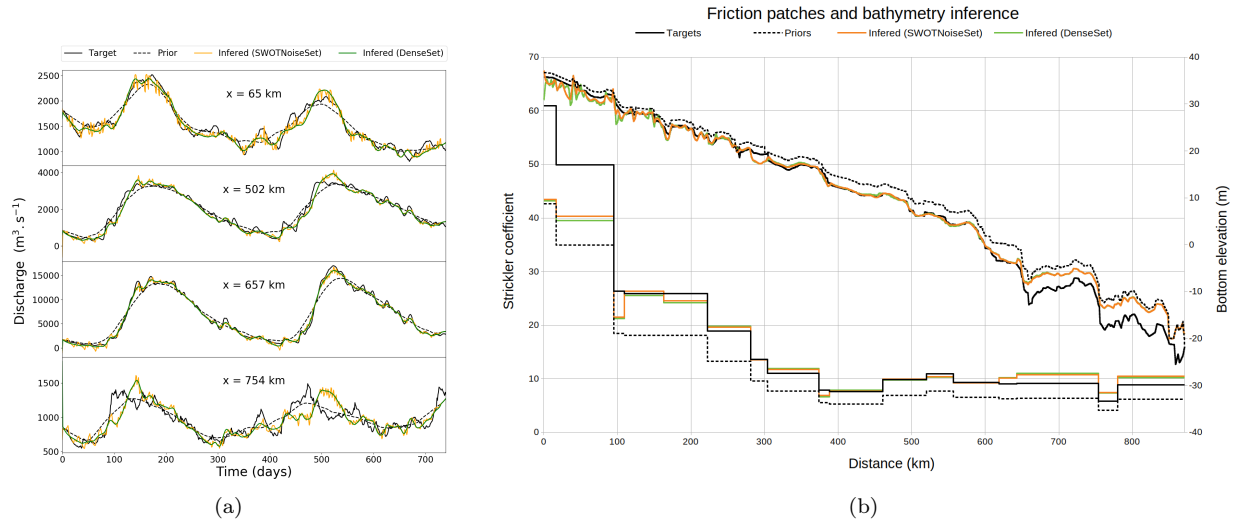


Figure 15: Inflow, bathymetry and friction patch inferences from SWOT synthetic data :  $c_{filter}^*$ , inferred control vector with a priori hydraulic behaviour.

Obs set	Prior	$RMSE [m^3 s^{-1}] (rRMSE)$				Nash-Sutcliffe			
		65 km	502 km	657 km	754 km	65 km	502 km	657 km	754 km
DenseSet	$Q_{FG}^{flat}$	189 (0.12)	329 (0.17)	1472 (0.22)	430 (0.45)	0.84	0.90	0.87	0.10
NoiseSWOTSet		209 (0.14)	360 (0.18)	1719 (0.26)	421 (0.44)	0.81	0.88	0.82	0.10
DenseSet	$Q_{FG}^{filter}$	101 (0.07)	195 (0.10)	412 (0.06)	158 (0.17)	0.94	0.97	0.99	0.49
NoiseSWOTSet		102 (0.07)	208 (0.11)	503 (0.07)	154 (0.16)	0.94	0.96	0.99	0.51

(a) Hydrograph scores

Obs set	Prior	$RMSE_{\alpha} [m^{1/3} s^{-1}]$	$RMSE_b [m]$
DenseSet	$Q_{FG}^{flat}$	5.35	1.89
NoiseSWOTSet		5.12	1.64
DenseSet	$Q_{FG}^{filter}$	5.30	2.06
NoiseSWOTSet		5.13	1.99

(b) Channel parameters scores

Table 6: Inferred parameter scores for extended control inferences.

of the anabranching Negro River in the Amazon basin using multisatellite data.

Through low Froude synthetic cases, it is shown that the signature of a lateral inflow is visible downstream from the inflow point through the total flow signature and can be visible upstream in case of downstream control at the injection. Following this analysis and using the HiVDI variational assimilation method (global in time and space), a study of the minimum spatial density of water surface observations necessary to infer lateral inflows from their hydraulic signatures is carried out. Synthetic twin experiments yield the following results: (i) given high observation temporal frequency relative to model hydraulic response, perfect inflows inferences can be obtained; (ii) to correctly attribute signatures between multiple lateral inflows, a minimum of 1 observation point between each injection cell is necessary; (iii) when simultaneously inferring inflows and/or channel parameters, a sensitivity to parameter weights (see section AppendixA) appears; (iv) given a priori parameter weights, accurate inferences of inflows and channel parameters is achievable even with the minimum spatial observability.

A method for building effective river models in coherence with multisatellite data and including realistic spatial variations is introduced based on multisource data of water surface elevation, width and slopes. This method makes use of (i) multitemission altimetric rating curves (see [64]) or equivalently a distributed hydrological model and altimetry data and (ii) water surface width like those from current databases (see [71]); it should be applicable to rivers from the future SWOT database. It is applied here to build a simple effective 1D model of the Negro river upstream from its confluence with the Solimoes river. It fits currently available satellite water surfaces signatures and contains real-like spatial variabilities and flood wave propagation features.

The inference capabilities of spatially distributed channel parameters and inflows from synthetic SWOT observations are highlighted on the Negro River case given hydraulically coherent priors. The inference of temporal parameters in the form of 21 spatially distributed lateral inflow hydrographs leads to accurate estimates and low

687 water surface misfit at observation times. High frequency observations give good inferences, with an expected  
688 sensitivity to both prior bias (see [1, 34]) and prior shape.

689 SWOT-like observations lead to comparable inferences, with slight oscillations due to the frequential disparity  
690 between observations temporal controls combined to their spatial distribution and the resulting nonlinear flow prop-  
691 agation on the domain, as analyzed with identifiability maps. Those oscillations related to model-observations time  
692 scales could be overcome by introducing additional regularizations - not done here for the sake of hydraulic analysis.  
693 Inference of purely spatial parameters (bathymetry/friction) were carried out as well, leading to some complemen-  
694 tary remarks: (i) channel parameters equifinality is most present in the downstream part of the model, where the  
695 waterline is influenced by the strong backwater effect of the boundary condition (Solimoes River) which diminishes  
696 water surface sensitivity to other parameters; (ii) bathymetry prior shape influences the inferred bathymetry. Fi-  
697 nally, simultaneous inference of channel parameters and spatially distributed inflows was achieved with satisfying  
698 accuracy. We show that, with the present method, large scale temporal parameter variations can be found from  
699 synthetic SWOT observations even without a priori knowledge of the shape of the hydrological response, but that  
700 small scale variations can be better inferred with a priori hydrograph shape knowledge.

701 Recall that the estimation of discharges and channel parameters from (SWOT) WS observations is a difficult  
702 inverse problem because of the correlated influence of flow controls on the observable water surface signatures -  
703 non uniqueness/equifinality issues. It is therefore necessary to use hydraulically consistent priors as investigated in  
704 [1, 37, 34] with HiVDI method that contains low complexity flow relations for deriving robust prior guesses from  
705 databases and hydrological models, or even in situ depth/discharge data when available (see [1]) - not the scope of  
706 this paper. As already discussed in [1], the VDA solution depends on the priors which are the first guess value and  
707 the covariances matrixes. Ongoing research efforts in the SWOT community, in view of global discharge estimates,  
708 focus on the determination of priors through the construction of a SWOT a-priori database based on [72, 71] and  
709 global/regional model outputs (see [26, 1]), constrained with available in situ gauge measurements. Note that  
710 a-priori estimations/databases could be enriched or reprocessed during or after the SWOT mission lifetime and  
711 HiVDI would enable to refining discharge estimates (see [1]). Moreover, as shown in [73], priors obtained by deep  
712 learning can greatly improve global estimation.

713 More generally, reaching unbiased estimates of discharge, from downstream to upstream of river networks with  
714 varying densities of in situ discharge data hence ungauged river portions/basins, is a crucial challenge in hydrology  
715 that could benefit from the fusion of complementary in situ and remotely sensed data in integrated models. The  
716 present study brings insight in lateral inflows inference from hydraulic signatures and paves the way for further  
717 research on integrated hydrological-hydraulic assimilation chains for river networks and in coherence with multi-  
718 satellites observables (of local hydrodynamic signatures) to benefit from them in a regionalization perspective.

719 Searching for distributed channel parameters and inflows given temporally sparse SWOT data and a global  
720 assimilation method brings the issue of signal attribution to the forefront, especially at the scale of a river network.

Further research should focus on tributaries that could be amenable to the use of SWOT and multisatellite observations to better constrain estimates of lateral inflows and next distributed fluxes on network models considering hydrological-hydraulic couplings. The introduction of pertinent signatures, scales and constraints in the forward-inverse models (e.g. forward operators, covariance matrices, cost function terms) should help maximizing the use of various information sources and enable smooth discharge estimates and better signal attribution, given uneven and heterogeneous satellite data in combination with other complementary databases/knowledge. This could help leveraging better inferences of hydrological responses and flow structure within a river basin and eventually enable information feedback to rainfall-runoff modules and ultimately regionalization issues.

## Author contributions and acknowledgments

This work is a part of the PhD thesis work of LP.

Research plan: LP, PAG, JM, PFG; computational software DassFlow1D and satellite data curation toolbox adapted from their previous versions by LP; Numerical investigations by LP with PAG, PFG, KL, JM for analysis. JM is the principal designer of the inverse computational method and its analysis.

Data : The water surface widths were derived by SB and HY contributed to their analysis using satellite images. The SWOT synthetic observations were generated by KL. The multisatellite dataset was provided by SC and altimetric rating curves and MGB outputs by AP. The GPS flow lines were provided by DM and SC.

Fundings: Most authors have been co-funded by CNES Tosca research project 2014-2019; project supervision: JM, PAG. PhD of LP is co-funded by CNES and ICUBE.

- [1] K. Larnier, J. Monnier, P-A. Garambois, and J. Verley. On the estimation of river discharges from altimetry. *Inverse Problems Sc. Eng. (IPSE)*. Accepted, to appear, 2020. URL <https://hal.archives-ouvertes.fr/hal-01811683>.
- [2] D.M. Moreira. *Apport des données de géodésie spatiale pour l'étude du bassin hydrologique amazonien*. PhD thesis, LEGOS - Laboratoire d'Etudes en Géophysique et Océanographie Spatiale, 2016. URL <http://www.theses.fr/2016TOU30372>.
- [3] J-F. Pekel, A. Cottam, and N. Gorelick. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(418):19, 2016. doi: 10.1038/nature20584. URL <https://www.nature.com/articles/nature20584#supplementary-information>.
- [4] R.C.D. Paiva, D.C. Buarque, W. Collischonn, M-P. Bonnet, F. Frappart, S. Calmant, and E.M. Carlos. Large-scale hydrologic and hydrodynamic modeling of the amazon river basin. *Water Resources Research*, 49(3): 1226–1243, 2013. doi: 10.1002/wrcr.20067. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20067>.

- [5] P. Brisset, J. Monnier, P-A. Garambois, and H. Roux. On the assimilation of altimetric data in 1D Saint-Venant river flow models. *Advances in water resources*, 119:41–59, 2018.
- [6] World Meteorological Organization. GCOS/GTOS plan for terrestrial climate-related observations, version 2.0. URL [https://library.wmo.int/doc\\_num.php?explnum\\_id=3891](https://library.wmo.int/doc_num.php?explnum_id=3891).
- [7] World Meteorological Organization. GCOS systematic observation requirements for satellite-based products for climate 2011 update: Supplemental details to the satellite-based component of the implementation plan for the global observing system for climate in support of the UNFCCC (2010 update)., 2011. URL [https://library.wmo.int/doc\\_num.php?explnum\\_id=3710](https://library.wmo.int/doc_num.php?explnum_id=3710).
- [8] J. Schuite, N. Flipo, N. Massei, A. Rivière, and F. Baratelli. Improving the spectral analysis of hydrological signals to efficiently constrain watershed properties. *Water Resources Research*, 55(5):4043–4065, 2019. doi: 10.1029/2018WR024579. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024579>.
- [9] B.M. Fekete and C.J. Vorosmarty. The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements. *IAHS - PUB*, 2002.
- [10] D. Alsdorf, P. Bates, J. Melack, M. Wilson, and T. Dunne. Spatial and temporal complexity of the Amazon flood measured from space. *Geophysical Research Letters*, 34(8), 2007. doi: 10.1029/2007GL029447. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL029447>.
- [11] M. Durand, L. Fu, D.P. Lettenmaier, D.E. Alsdorf, E. Rodríguez, and D. Esteban-Fernandez. The Surface Water and Ocean Topography Mission: Observing Terrestrial Surface Water and Oceanic Submesoscale Eddies. *Proceedings of the IEEE*, 98(5):766–779, May 2010. doi: 10.1109/JPROC.2010.2043031.
- [12] E. Rodríguez. SWOT Science requirements document. JPL document, JPL, 2012. URL [https://swot.jpl.nasa.gov/system/documents/files/2176\\_2176\\_D-61923\\_SRD\\_Rev\\_B\\_20181113.pdf](https://swot.jpl.nasa.gov/system/documents/files/2176_2176_D-61923_SRD_Rev_B_20181113.pdf).
- [13] S. Biancamaria, D.P. Lettenmaier, and T.M. Pavelsky. The SWOT mission and its capabilities for land hydrology. *Surveys in Geophysics*, 37(2):307–337, Mar 2016. ISSN 1573-0956. URL <http://dx.doi.org/10.1007/s10712-015-9346-y>.
- [14] E. Rodríguez, Esteban-Fernandez D., E. Peral, C.W. Chen, J-W. Blesser, and B. Williams. Wide-swath altimetry: A review. in D. Stammer A. Cazenave (Eds.). *Satellite Altimetry over Oceans and Land Surfaces (Chapter 2)*, page CRC press, 2018.
- [15] P-A. Garambois, S. Calmant, H. Roux, A. Paris, J. Monnier, P. Finaud-Guyot, A.S. Montazem, and J.S. da Silva. Hydraulic visibility: Using satellite altimetry to parameterize a hydraulic model of an ungauged

reach of a braided river. *Hydrological Processes*, 31(4):756–767, 2017. ISSN 1099-1085. URL <http://dx.doi.org/10.1002/hyp.11033>. hyp.11033.

[16] A.S. Montazem, P-A. Garambois, S. Calmant, P. Finaud-Guyot, J. Monnier, D.M. Moreira, J.T. Minear, and S. Biancamaria. Wavelet-based river segmentation using hydraulic control-preserving water surface elevation profile properties. *Geophysical Research Letters*, 46(12):6534–6543, JUN 28 2019. ISSN 0094-8276. doi: 10.1029/2019GL082986.

[17] E. Rodríguez, M. Durand, and R.P. de M. Frasson. Observing rivers with varying spatial scales. *Water Resources Research*, 2020. doi: 10.1029/2019WR026476. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026476>.

[18] P-A. Garambois and J. Monnier. Inference of effective river properties from remotely sensed observations of water surface. *Advances in Water Resources*, 79:103–120, 2015. URL <https://www.sciencedirect.com/science/article/abs/pii/S0309170815000330>.

[19] K.M. Andreadis, E.A. Clark, D.P. Lettenmaier, and D.E. Alsdorf. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34(10), 2007. doi: 10.1029/2007GL029721. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL029721>.

[20] S. Biancamaria, M. Durand, K.M. Andreadis, P.D. Bates, A. Boone, N.M. Mognard, E. Rodríguez, D.E. Alsdorf, D.P. Lettenmaier, and E.A. Clark. Assimilation of virtual wide swath altimetry to improve arctic river modeling. *Remote Sensing of Environment*, 115(2):373 – 381, 2011. ISSN 0034-4257. doi: 10.1016/j.rse.2010.09.008. URL <http://www.sciencedirect.com/science/article/pii/S0034425710002816>.

[21] M. Durand, K.M. Andreadis, D.E. Alsdorf, D.P. Lettenmaier, D. Moller, and M. Wilson. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, 35(20), 2008. doi: 10.1029/2008GL034150. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GL034150>.

[22] Y. Yoon, M. Durand, C.J. Merry, E.A. Clark, K.M. Andreadis, and Alsdorf D.E. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, 464 - 465(0):363 – 375, 2012. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2012.07.028. URL <http://www.sciencedirect.com/science/article/pii/S0022169412006294>.

[23] M. Durand, J. Neal, E. Rodríguez, K.M. Andreadis, L.C. Smith, and Y. Yoon. Estimating reach-averaged discharge for the river Severn from measurements of river water surface elevation and slope. *Journal of Hydrology*, 511:92–104, 2014. doi: 10.1016/j.jhydrol.2013.12.050.

- [24] C.J. Gleason and L.C. Smith. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, 111(13):4788–4791, 2014. ISSN 0027-8424. doi: 10.1073/pnas.1317606111. URL <http://www.pnas.org/content/111/13/4788>.
- [25] D.M. Bjerklie, C.M. Birkett, .W. Jones, C. Carabajal, J.A. Rover, J.W. Fulton, and P-A. Garambois. Satellite remote sensing estimation of river discharge: Application to the Yukon river Alaska. *Journal of Hydrology*, 561: 1000 – 1018, 2018. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2018.04.005. URL <http://www.sciencedirect.com/science/article/pii/S0022169418302464>.
- [26] M. Durand, C.J. Gleason, P-A. Garambois, D. Bjerklie, L.C. Smith, H. Roux, E. Rodriguez, P.D. Bates, T.M. Pavelsky, J. Monnier, X. Chen, G. Di Baldassarre, J-M. Fiset, N. Flipo, R.P.d.M. Frasson, J. Fulton, N. Goutal, F. Hossain, E. Humphries, J.T. Minear, M.M. Mukolwe, J.C. Neal, S. Ricci, B.F. Sanders, G. Schumann, J.E. Schubert, and L. Vilmin. An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, 52(6):4527–4549, 2016. doi: 10.1002/2015WR018434. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR018434>.
- [27] H. Roux and D. Dartus. Use of parameter optimization to estimate a flood wave: Potential applications to remote sensing of rivers. *J. of Hydrology*, 328:258–266, 2006.
- [28] M. Honnorat, X. Lai, F-X. le Dimet, and J. Monnier. Variational data assimilation for 2D fluvial hydraulics simulation. *CMWR XVI-Computational Methods for Water Ressources. Copenhagen, june 2006*.
- [29] R. Hostache, X. Lai, J Monnier, and C. Puech. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. *Journal of Hydrology*, 390:257–268, 2010. URL <http://www.sciencedirect.com/science/article/pii/S0022169410004166>.
- [30] X. Lai and J. Monnier. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: mathematical method and test case. *Journal of Hydrology*, 377:1–11, 2009. URL <http://www.sciencedirect.com/science/article/pii/S0022169409004508>. 1-2.
- [31] D.G. Cacuci, I.M. Navon, and M. Ionescu-Bugor. *Computational Methods for Data Evaluation and Assimilation*. Taylor and Francis CRC Press: Boca Raton, 2013.
- [32] I. Gejadze and P-O. Malaterre. Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model. *International Journal for Numerical Methods in Fluids*, 83(5):405–430, 2017. ISSN 1097-0363. doi: 10.1002/flid.4273. URL <http://dx.doi.org/10.1002/flid.4273>.
- [33] H. Oubanas, I. Gejadze, P-O. Malaterre, and F. Mercier. River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model. *Journal of*



*Hydrology*, 559:638 – 647, 2018. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2018.02.004. URL <http://www.sciencedirect.com/science/article/pii/S0022169418300805>.

[34] P-A. Garambois, K. Larnier, J. Monnier, P. Finaud-Guyot, J. Verley, A.S. Montazem, and S. Calmant. Variational estimation of effective channel and ungauged anabranching river discharge from multi-satellite water heights of different spatial sparsity. *Journal of Hydrology*, 581:124409, 2020. ISSN 0022-1694. URL <https://doi.org/10.1016/j.jhydrol.2019.124409>.

[35] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems*, volume 6. Walter de Gruyter, 2008.

[36] Y. Yoon, P-A. Garambois, R. Paiva, M. Durand, H. Roux, and E. Beighley. Improved error estimates of a discharge algorithm for remotely sensed river measurements: Test cases on Sacramento and Garonne rivers. *Water Resources Research*, 52(1):278–294, 2016.

[37] S. Tuozzolo, G. Lind, B. Overstreet, J. Mangano, M. Fonstad, M. Hagemann, R.P.M. Frasson, K. Larnier, P-A. Garambois, J. Monnier, and M. Durand. Estimating river discharge with swath altimetry: A proof of concept using AirSWOT observations. *Geophysical Research Letters*, 46(3):1459–1466, 2019. doi: 10.1029/2018GL080771. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL080771>.

[38] H. Oubanas, I. Gejadze, P-O. Malaterre, M. Durand, R. Wei, R.P.M. Frasson, and A. Domeneghetti. Discharge estimation in ungauged basins through variational data assimilation: The potential of the SWOT mission. *Water Resources Research*, 54(3):2405–2423, 2018. doi: 10.1002/2017WR021735. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR021735>.

[39] R. Schneider, P.N. Godiksen, H. Villadsen, H. Madsen, and P. Bauer-Gottwein. Application of CryoSat-2 altimetry data for river analysis and modelling. *Hydrology and Earth System Sciences*, 21(2):751–764, 2017. doi: 10.5194/hess-21-751-2017. URL <https://www.hydrol-earth-syst-sci.net/21/751/2017/>.

[40] F.E. O’Loughlin, J. Neal, G.J.P. Schumann, E. Beighley, and P.D. Bates. A LISFLOOD-FP hydraulic model of the middle reach of the Congo. *Journal of Hydrology*, page 124203, 2019. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2019.124203. URL <http://www.sciencedirect.com/science/article/pii/S0022169419309382>.

[41] W. Collischonn, D. Allasia, B.C. Da Silva, and E.C.M. Tucci. The MGB-IPH model for large-scale rainfall-runoff modelling. *Hydrological Sciences Journal*, 52(5):878–895, 2007. URL <https://doi.org/10.1623/hysj.52.5.878>.

[42] P.R.M. Pontes, F.M. Fan, A.S. Fleischmann, R.C.D. Paiva, D.C. Buarque, V.A. Siqueira, P.F. Jardim, M.V. Sorribas, and W. Collischonn. MGB-IPH model for hydrological and hydraulic simulation of large floodplain river systems coupled with open source GIS. *Environmental Modelling Software*, 94:1 – 20, 2017. ISSN

1364-8152. doi: 10.1016/j.envsoft.2017.03.029. URL <http://www.sciencedirect.com/science/article/pii/S136481521630189X>.

[43] R.C.D. Paiva, D.C. Buarque, W. Collischonn, M-P. Bonnet, F. Frappart, S. Calmant, and E.M. Carlos. Large-scale hydrologic and hydrodynamic modeling of the amazon river basin. *Water Resources Research*, 49(3): 1226–1243, 2013. doi: 10.1002/wrcr.20067. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20067>.

[44] Adhemar J-C. de Saint-Venant. Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l’introduction des marées dans leur lit. *CR Acad. Sci. Paris*, 73(147-154):5, 1871.

[45] V.T. Chow. Handbook of applied hydrology. *McGraw-Hill Book Co., New-York*, 1467 pages, 1964.

[46] C. Ancey. *Hydraulique à surface libre*. Ecole Polytechnique Fédérale de Lausanne Ecublens, 2018.

[47] V. Guinot. *Wave propagation in fluids: models and numerical techniques, second edition.*, volume 49. 2010.

[48] J. A. Cunge, M. Holly, F., and A. Verwey. *Practical Aspects of Computational River Hydraulics*. Pitam Publishing,, 1980.

[49] Data assimilation for free surface flows. Technical report, Mathematics Institute of Toulouse-INSA group-CS corp.CNES-CNRS, 2019. URL <http://www.math.univ-toulouse.fr/DassFlow>.

[50] J. Monnier, F. Couderc, D. Dartus, K. Larnier, R. Madec, and J-P. Vila. Inverse algorithms for 2D shallow water equations in presence of wet dry fronts. application to flood plain dynamics. *Advances in Water Resources*, 97:11–24, 2016.

[51] J. Monnier. *Variational data assimilation: from optimal control to large scale data assimilation*. Open On-line Course, INSA Toulouse, 2014. URL <https://www.math.univ-toulouse.fr/%7Ejmonnie/Enseignement/CourseVDA.pdf>.

[52] A. Montazem, P.-A. Garambois, P. Finaud-Guyot, S. Calmant, J. Monnier, and D. Moreira. Physical basis for river segmentation from water surface observables. In *AGU Fall Meeting Abstracts*, volume 2017, pages H44H–06, December 2017. URL <https://ui.adsabs.harvard.edu/abs/2017AGUFM.H44H..06S>.

[53] P.G. Samuels. Backwater lengths in rivers. *Proceedings of the Institution of Civil Engineers*, 87(4):571–582, 1989. URL <https://doi.org/10.1680/iicep.1989.3779>.

[54] V.T. Chow. *Open-channel Hydraulics*. Mc Graw-Hill, New-York, USA, 1959.

[55] A. Montazem. *Représentation et segmentation hydraulique effective de rivières pour le calcul de debit par altimetrie SWOT à l’échelle globale*. Thèse de doctorat, Universite de Toulouse III Paul Sabatier, Toulouse, France, 2018.

- [56] Agência Nacional de Águas (ANA). Brasil hidrografia. [http://www.portalbrasil.net/brasil\\_hidrografia.htm](http://www.portalbrasil.net/brasil_hidrografia.htm). Accessed: 2020-02-06.
- [57] E.M. Latrubesse and E. Franzinelli. The late quaternary evolution of the negro river, amazon, brazil: Implications for island and floodplain formation in large anabranching tropical systems. *Geomorphology*, 70(3):372 – 397, 2005. ISSN 0169-555X. doi: 10.1016/j.geomorph.2005.02.014. URL <http://www.sciencedirect.com/science/article/pii/S0169555X05000826>. Tropical Rivers.
- [58] G. Cochonneau, F. Sondag, J-L. Guyot, B. Geraldo, N. Filizola, P. Fraizy, A. Laraque, P. Magat, J-M. Martinez, L. Noriega, E. Oliveira, J. Ordonez, R. Pombosa, Seyler F., J. Sidgwick, and P. Vauchel. *L’Observatoire de Recherche en Environnement, ORE HYBAM sur les grands fleuves amazoniens*. 2006. ISBN 978-1-901502-78-7.
- [59] N. Filizola, N. Spínola, W. Arruda, Fr. Seyler, S. Calmant, and J. Silva. The Rio Negro and Rio Solimões confluence point-hydrometric observations during the 2006/2007 cycle. *River, Coastal and Estuarine Morphodynamics-RCEM*, 2009:1003–1006, 2009.
- [60] J. Callède, D.M. Moreira, and S. Calmant. Détermination de l’altitude du zéro des stations hydrométriques en Amazonie brésilienne. Application aux lignes d’eau des Rios Negro, Solimões et Amazone. *Journal of Water Science*, 26(2):153–171, 2013. doi: 10.7202/1016065ar.
- [61] J.S. Da Silva, F. Seyler, S. Calmant, O.C. Rotunno Filho, E. Roux, A.A.M. Arango, and J-L. Guyot. Water level dynamics of amazon wetlands at the watershed scale by satellite altimetry. *International Journal of Remote Sensing*, 33(11):3323–3353, 2012. URL <http://dx.doi.org/10.1080/01431161.2010.531914>.
- [62] G.H. Allen and T.M. Pavelsky. Patterns of river width and surface area revealed by the satellite-derived north american river width data set. *Geophysical Research Letters*, 42(2):395–402, 2015. ISSN 1944-8007. URL <http://dx.doi.org/10.1002/2014GL062764>.
- [63] E. Park and E.M. Latrubesse. The hydro-geomorphologic complexity of the lower amazon river floodplain and hydrological connectivity assessed by remote sensing and field control. *Remote Sensing of Environment*, 198: 321 – 332, 2017. ISSN 0034-4257. doi: 10.1016/j.rse.2017.06.021. URL <http://www.sciencedirect.com/science/article/pii/S0034425717302808>.
- [64] A. Paris, R.C.D. de Paiva, J.S. da Silva, D.M. Moreira, S. Calmant, P-A. Garambois, W. Collischonn, M-P. Bonnet, and F. Seyler. Stage-discharge rating curves based on satellite altimetry and modeled discharge in the amazon basin. *Water Resources Research*, 52(5):3787–3814, 2016. doi: 10.1002/2014WR016618. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016618>.
- [65] A.S. Fleischmann, W. Collischonn, R. Paiva, and C.E. Tucci. Modeling the role of reservoirs versus floodplains

on large-scale river hydrodynamics. *Natural Hazards*, 99(2):1075–1104, Nov 2019. ISSN 1573-0840. URL <https://doi.org/10.1007/s11069-019-03797-9>.

[66] A.C. Fassoni-Andrade, R.C.D. de Paiva, and A.S. Fleischmann. Lake topography and active storage from satellite observations of flood frequency. *Water Resources Research*, 56(7), 2020. doi: 10.1029/2019WR026362. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019WR026362>.

[67] A.S. Fleischmann, R.C.D. Paiva, W. Collischonn, M.V. Sorribas, and P.R.M. Pontes. On river-floodplain interaction and hydrograph skewness. *Water Resources Research*, 52(10):7615–7630, 2016. doi: 10.1002/2016WR019233. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016WR019233>.

[68] W. Collischonn, A.S. Fleischmann, R.C.D. Paiva, and A. Mejia. Hydraulic causes for basin hydrograph skewness. *Water Resources Research*, 53(12):10603–10618, 2017. doi: 10.1002/2017WR021543. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR021543>.

[69] O. Thual. *Hydrodynamique de l’environnement*. Editions de l’Ecole Polytechnique, 2010.

[70] CNES. SWOT hydrology toolbox. URL <https://github.com/CNES/swot-hydrology-toolbox>.

[71] G.H. Allen and T.M. Pavelsky. Global extent of rivers and streams. *Science*, 2018. ISSN 0036-8075. doi: 10.1126/science.aat0636. URL <http://science.sciencemag.org/content/early/2018/06/27/science.aat0636>.

[72] Konstantinos M. Andreadis, Guy J-P. Schumann, and Tamlin Pavelsky. A simple global river bankfull width and depth database. *Water Resources Research*, 49(10):7164–7168, 2013. ISSN 1944-7973. URL <http://dx.doi.org/10.1002/wrcr.20440>.

[73] K. Larnier and J. Monnier. Hybrid neural network - data assimilation algorithm to infer river features. *submitted*, 2020.

[74] J. C. Gilbert and C. Lemaréchal. Some numerical experiments with variable-storage quasi-newton algorithms. *Mathematical programming*, 45(1-3):407–435, 1989.

[75] L. Hascoët and V. Pascual. The Tapenade Automatic Differentiation tool: Principles, Model, and Specification. *ACM Transactions On Mathematical Software*, 39(3), 2013. URL <http://dx.doi.org/10.1145/2450153.2450158>.

[76] F. Bouttier and P. Courtier. Data assimilation concepts and methods march 1999. *Meteorological training course lecture series. ECMWF*, page 59, 2002. URL [http://msi.ttu.ee/~elken/Assim\\_concepts.pdf](http://msi.ttu.ee/~elken/Assim_concepts.pdf).

- [77] A.C. Lorenc, S.P. Ballard, R.S. Bell, N.B. Ingleby, P.L.F. Andrews, D.M. Barker, J.R. Bray, A.M. Clayton, T. Dalby, D. Li, et al. The met. office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000.
- [78] A. Weaver and P. Courtier. Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 127(575):1815–1846, 2001.
- [79] S.A. Haben, A.S. Lawless, and N.K. Nichols. Conditioning and preconditioning of the variational data assimilation problem. *Computers & Fluids*, 46(1):252–256, 2011.
- [80] S.A. Haben, A.S. Lawless, and N.K. Nichols. Conditioning of incremental variational data assimilation, with application to the met office system. *Tellus A*, 63(4):782–792, 2011.

## AppendixA. The computational inverse method

The computational inverse method is based on Variational Data Assimilation (VDA) applied to the 1D Saint-Venant model (eq. (1)). The computational inverse method is the one presented in [5, 1] with an augmented composite control vector  $c$  (eq. (4)):  $c$  contains a spatially distributed friction coefficient enabling to model complex flow zones (while it is an uniform friction law  $K(h)$  in [1]). This definition of  $K(x, h)$  enables to consider more heterogeneous bathymetry controls.  $c$  also contains lateral flow hydrographs  $Q_{lat,i}(t)$  to deal with in/offtakes due to tributaries or underground flows. It is important to point out that the imposed downstream BC is an unknown of the inverse problem. It is constrained with the observed water elevations and inferred river bottom slope using a locally uniform flow hypothesis (i.e. Manning equation, see subsection 2.1) - except in the last real case above. The cost function  $j(c)$  is defined as:

$$j(c) = j_{obs}(c) + \gamma j_{reg}(c) \quad (\text{A.1})$$

where  $\gamma > 0$  is a weighting coefficient of the so-called “regularization term”  $j_{reg}(c)$ . The term  $j_{obs}(c)$  measures the misfit between observed and modeled WS elevations such that:

$$j_{obs}(c) = \frac{1}{2} \| (Z(c) - Z_{obs}) \|_{\mathcal{O}}^2 \quad (\text{A.2})$$

The norm  $\| \cdot \|_{\mathcal{O}} = \| \mathcal{O}^{1/2} \cdot \|_2$  is defined from an a priori positive definite covariance matrix  $\mathcal{O}$ . Assuming uncorrelated observations  $\mathcal{O} = \text{diag}(\sigma_Z)$ . The modeled WS elevations  $Z$  depend on  $c$  through the hydrodynamic model (eq. (1)) and the inverse problem reads as

$$c^* = \text{argmin}_c j(c) \quad (\text{A.3})$$

990 This optimal control problem is solved using a Quasi-Newton descent algorithm: the L-BFGS algorithm version  
 991 presented in [74]. The cost gradient  $\nabla j(c)$  is computed by solving the adjoint model; the latter is obtained by  
 992 automatic differentiation using Tapenade software [75]. Detailed know-hows on VDA may be found e.g. in the  
 993 online courses [76, 51].

994 To be solved efficiently this optimization problem needs to be “regularized”. Indeed the friction and the  
 995 bathymetry may trigger indiscernible surface signatures therefore leading to an ill-posed inverse problem; we refer  
 996 e.g. to [35] for the theory of regularization of such inverse problems and to [1] for a discussion focused on the present  
 997 inverse flow problem.

998 Following [1], the optimization problem (eq. (A.3)) is regularized as follows. First the regularization term  $j_{reg}$   
 999 is added to the cost function (eq. (A.1)). We simply set:  $j_{reg}(c) = \frac{1}{2} \|b''(x)\|_2^2$ . Therefore this term imposes (as  
 1000 weak constraints) the inferred bathymetry profile  $b(x)$  to be an elastic interpolating the values of  $b$  at the control  
 1001 points (i.e. a cubic spline).

1002 A specificity of the present context is the large inconsistency between the large observation grid (altimetry  
 1003 points) and the finer model grid. Between the sparse observations points (equivalently the control points), the  
 1004 bathymetry profile  $b(x)$  is reconstructed as a piecewise linear function. It is worth to point out that the resulting  
 1005 reconstruction is consistent with the physical analysis presented in [52, 55]. (This study analyses the adequation  
 1006 between the SW model (eq. (1)) behavior and the WS signature).

1007 Next and following [77, 78, 1], the following change of control variable is made:

$$k = B^{-1/2} (c - c_{prior}) \quad (\text{A.4})$$

1008 where  $c$  is the original control vector,  $c_{prior}$  is a prior value of  $c$  and  $B$  is a covariance matrix. The choice of  $B$  is  
 1009 crucial in the VDA formulation; its expression is detailed below. After this change of variable the new optimization  
 1010 problem reads:

$$\min_k J(k) \text{ with } J(k) = j(c) \quad (\text{A.5})$$

1011 It is easy to show that this leads to the following new optimality condition:  $B^{1/2} \nabla j(c) = 0$ ; somehow a  
 1012 preconditioned optimality condition. For more details and explanations we refer to 79, 80 and [1] in the present  
 1013 inversion context.

1014 Assuming uncorrelated controls the matrix  $B$  is defined as block diagonal such that  $B = \text{diag} (B_Q, B_{Qlat,1}, \dots, B_{Qlat,L}, B_b, B_\alpha, \dots)$   
 1015 Still following [1], the covariance matrices  $B_Q$ ,  $B_{Qlat}$  and  $B_b$  are set as the classical second order auto-regressive  
 1016 correlation matrices:

$$(B_Q)_{i,j} = (\sigma_Q)^2 \exp\left(-\frac{|t_j - t_i|}{\Delta t_Q}\right), (B_{Q_{lat,l}})_{i,j} = (\sigma_{Q_{lat}})^2 \exp\left(-\frac{|t_j - t_i|}{\Delta t_Q}\right) \text{ and } (B_b)_{i,j} = (\sigma_b)^2 \exp\left(-\frac{|x_j - x_i|}{L_b}\right) \quad (\text{A.6})$$

1017 The VDA parameters  $\Delta t_Q$  and  $L_b$  represent prior hydraulic scales and act as correlation lengths. We refer  
 1018 to [5] for a thorough analysis of the discharge inference in terms of frequencies and wave lengths and [1] in the  
 1019 present river-observation context. In the present study, the friction parameters applied to deca-kilometric patches  
 1020 are assumed to be uncorrelated thus the matrices  $B_\alpha$  and  $B_\beta$  are diagonal:

$$(B_\alpha)_{i,i} = (\sigma_\alpha)^2, (B_\beta)_{i,i} = (\sigma_\beta)^2 \quad (\text{A.7})$$

1021 The scalar values  $\sigma_\square$  may be viewed as variances ; their values are given in the numerical results section.

1022 Finally, in a noised observation context and to avoid overfitting noisy data, we denote by  $\delta$  the noise level such  
 1023 that  $\|Z_{obs} - Z_{true}\|^2 \leq \delta$  with  $Z_{obs}$  the observed and  $Z_{true}$  the true WS elevation profiles. A common technique  
 1024 to avoid overfitting noisy data, in the context of Tykhonov's regularization of ill-posed problems, is Morozov's  
 1025 discrepancy principle, (see e.g. [35] and references therein): the regularization parameter  $\gamma$  (see eq. (A.1)) is chosen  
 1026 a posteriori such that  $j$  does not decrease below the noise level.

# 1028 AppendixB. Extended friction calibration results

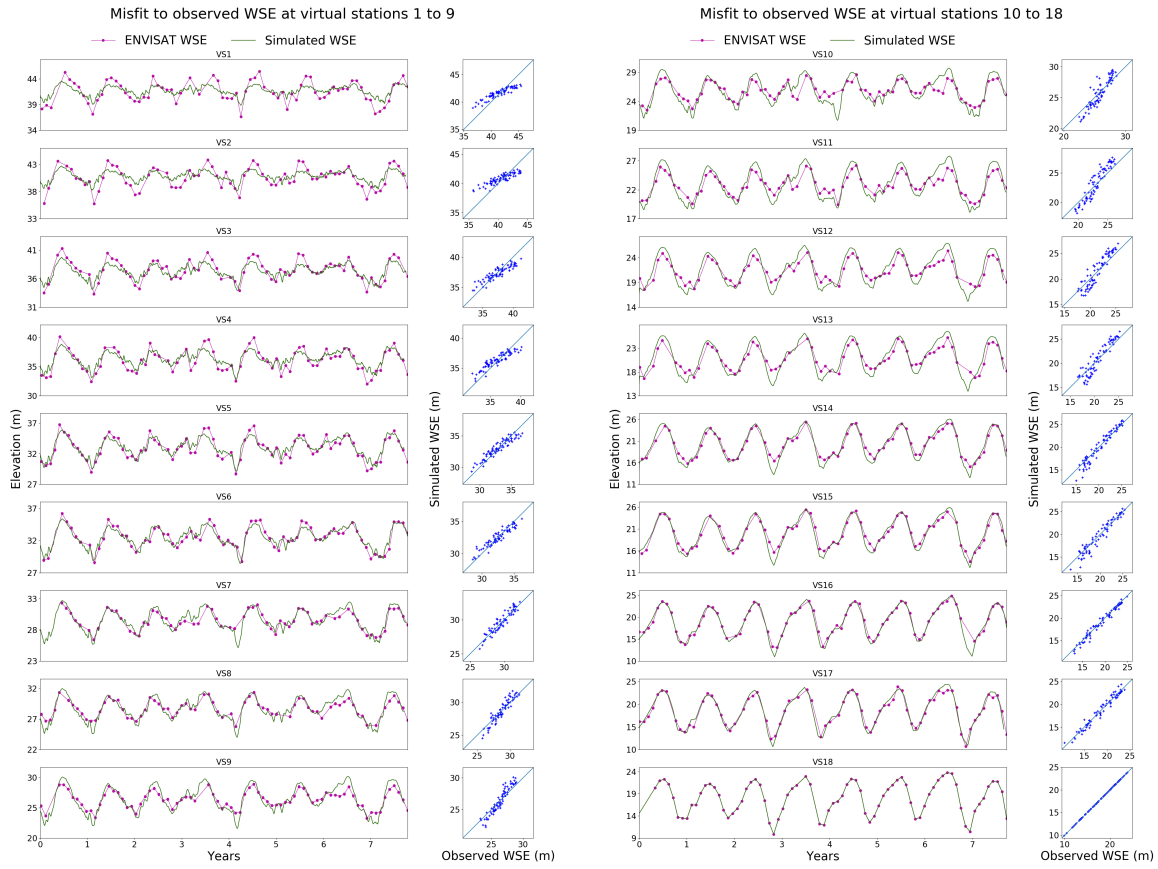


Figure B.16: ENVISAT WS elevation misfit after friction calibration at all stations for  $M2a$ .



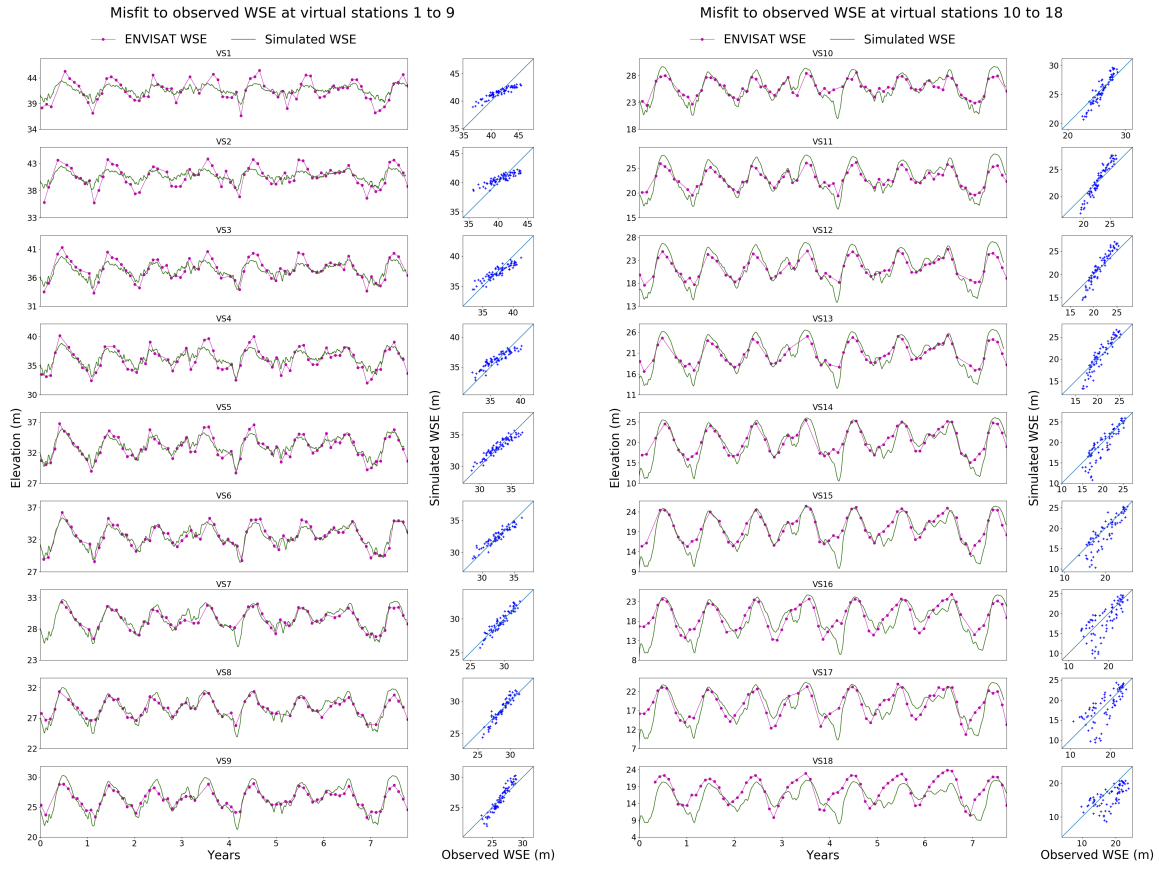


Figure B.17: ENVISAT WS elevation misfit after friction calibration at all stations for  $M2b$ .

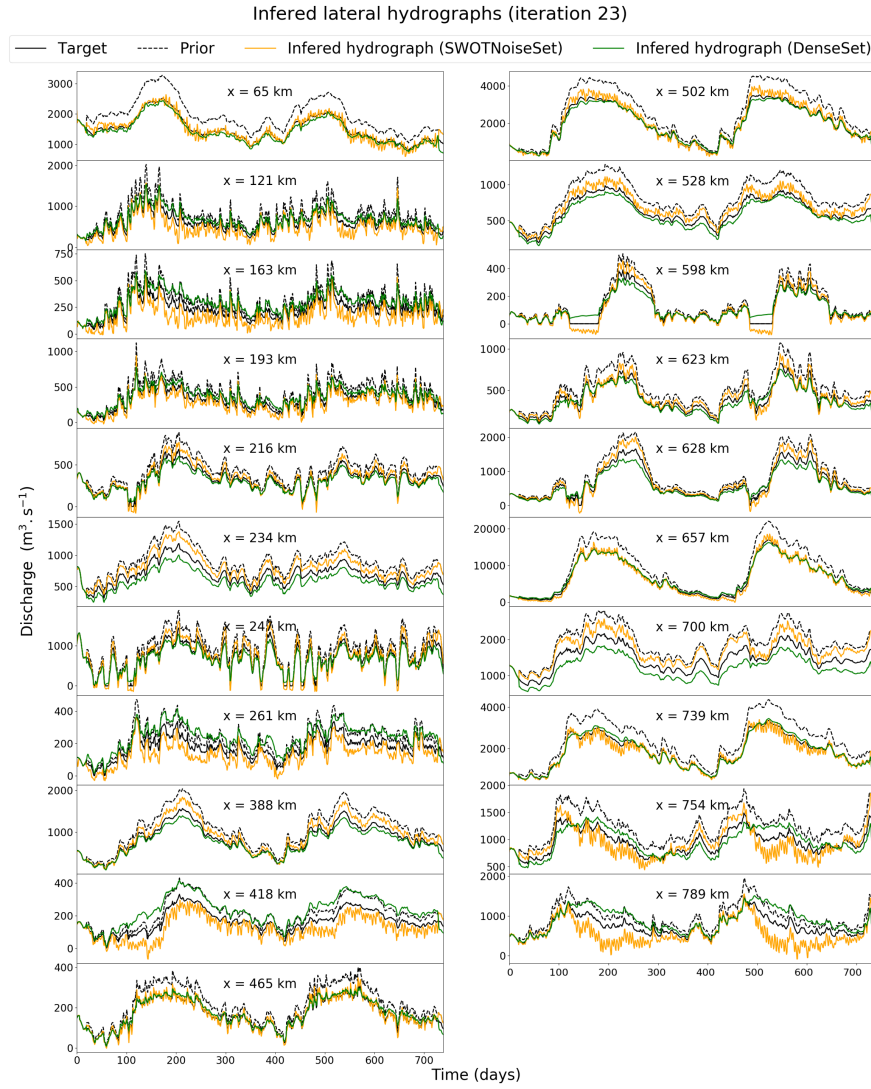


Figure C.18: Lateral hydrograph inferences from SWOTNoiseSet and DenseSet, using the  $Q_{FG}^{+30}$  inflow prior

$x$ (km)	$\mu$ (d)	$\sigma$ (m <sup>3</sup> s <sup>-1</sup> )	DenseSet RMSE (m <sup>3</sup> s <sup>-1</sup> )	SWOTNoiseSet RMSE (m <sup>3</sup> s <sup>-1</sup> )
65	1	2500	79.62	122.34
121		1500	97.74	134.59
163		1000	64.17	70.44
193		1000	49.06	50.75
216		700	37.04	41.48
234		800	112.76	120.43
247		1000	89.82	105.24
261		800	56.57	45.62
388		1000	98.30	125.94
418		800	60.64	49.12
465		600	8.86	12.64
502		2000	116.21	194.86
528		1000	66.85	83.55
598		600	28.69	29.08
623		1000	40.75	44.83
628		1200	119.85	126.13
657		5000	421.26	578.31
700		1500	263.07	264.27
739		2800	127.97	173.35
754		2000	140.89	158.19
789		2400	249.73	283.92

(a) With prior  $Q_{FG}^{-30}$

$x$ (km)	$\mu$ (d)	$\sigma$ (m <sup>3</sup> s <sup>-1</sup> )	DenseSet RMSE (m <sup>3</sup> s <sup>-1</sup> )	SWOTNoiseSet RMSE (m <sup>3</sup> s <sup>-1</sup> )
65	1	2500	134.66	156.34
121		1500	142.70	138.31
163		1000	69.46	66.83
193		1000	45.66	43.47
216		700	45.31	47.28
234		800	128.18	130.98
247		1000	122.21	125.72
261		800	35.82	33.85
388		1000	146.13	149.39
418		800	33.81	32.22
465		600	13.53	15.37
502		2000	212.87	235.58
528		1000	87.67	91.96
598		600	28.62	28.85
623		1000	39.51	43.31
628		1200	123.45	126.71
657		5000	581.03	775.15
700		1500	250.54	256.81
739		2800	157.54	194.14
754		2000	147.88	154.01
789		2400	293.07	297.37

(b) With prior  $Q_{FG}^{+30}$

Table C.7: Inferred lateral inflows parameter weights and RMSE

1033 **AppendixD. Technical specifications**

1034 **Host bridge:** Intel Corporation 8th Gen Core Processor Host Bridge/DRAM Registers

1035 **PCI bridge:** Intel Corporation Xeon E3-1200 v5/E3-1500 v5/6th Gen Core Processor PCIe Controller

1036 **Memory:** 2x16Gb SODIMM DDR4 Synchronous 2667 MHz (0.4 ns)

1037 **Resolution mode:** sequential

1038 **Resolution method:** implicit-explicit preissmann scheme

1039 **Sample run:** inference of the full triplet on the Negro model (inferred control vector  $c_{filter}^*$  in subsection 5.2)

- 1040     • Control vector components: 4x740 flow points, 436 bathymetry points, 17 friction patches (3413 total sought  
1041       values)
- 1042     • Total run time (direct): under 15 min
- 1043     • Total run time (inverse): 20.8 h
- 1044     • Number of iterations: 35
- 1045     • Average iteration time length: 35.8 min